

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 5, 2003, 06:17:30 ; Search time 58 Seconds
(without alignments)
863.552 Million cell updates/sec

Title: US-09-884-566A-2
Perfect score: 2650
Sequence: 1 MEVLESQSVLQWDRKLS.....LEGNETLKVLELRRVNAATP 521

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues
al number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 73:*
1: PIR1.*
2: PIR2.*
3: PIR3.*
4: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	445.5	16.8	515	2 A42140	box B-binding factor
2	436	16.5	516	2 A44494	CAMP-responsive el
3	303.5	11.5	688	2 T32750	hypothetical prote
4	298	11.2	365	2 T15656	hypothetical prote
5	252	9.5	699	2 T09069	probable CAMP-resp
6	167.5	6.3	741	2 I48694	probable transcrip
7	166	6.3	742	2 A49672	transcription fact
8	163	6.2	609	2 T49282	hypothetical prote
9	162.5	6.1	341	2 S26686	CAMP response elem
10	162.5	6.1	486	2 J42028	activating transcr
11	162	6.1	313	1 TVQJUN	transforming prote
12	161.5	6.1	278	2 S36101	CAMP response elem
13	160	6.0	322	2 T08592	TGACG-motif-bindin
14	159.5	6.0	296	1 TVFVJN	transforming prote
15	158.5	6.0	229	2 A37944	CAMP response elem
16	158.5	6.0	271	2 S12560	transcription fact
17	158	6.0	344	2 J55601	CAMP response elem
18	157.5	5.9	269	2 A46490	TCR V beta CRE-mot
19	157	5.9	772	2 A55004	transcription fact
20	156.5	5.9	310	2 I50373	jun protein - chic
21	155	5.8	322	2 T12093	TGACG-motif bindin
22	154.5	5.8	373	2 S32537	erythroid transcri
23	152.5	5.8	373	2 A49671	gene NF-E2 protein
24	152	5.7	334	1 TVMSJA	transcription fact
25	151.5	5.7	489	2 S50548	hypothetical prote
26	151.5	5.7	373	2 A54692	transcription fact
27	151.5	5.7	483	2 S32741	transcription fact
28	150	5.7	327	2 S22298	CAMP response elem
29	150	5.7	341	1 TVMSJD	transforming prote

30	150	5.7	341	2 JC4051	jun-D protein - ra
31	150	5.7	589	2 I59340	transcription fact
32	150	5.7	679	2 S48437	hypothetical prote
33	149.5	5.6	325	2 S23007	CAMP response elem
34	149.5	5.6	1844	1 RRWPTM	genome polypeptid
35	149	5.6	334	2 S12742	transcription fact
36	149	5.6	448	2 A42026	CAMP response elem
37	149	5.6	456	2 B42026	cyclic AMP respons
38	148.5	5.6	326	2 T08591	TGACG-motif bindin
39	148.5	5.6	328	2 A35663	CAMP response elem
40	148	5.6	360	2 T03373	probable G-box bin
41	148	5.6	372	2 T10472	G-box binding prot
42	147.5	5.6	341	2 S20827	CAMP response elem
43	147.5	5.6	505	1 S05380	transcription fact
44	147	5.5	326	2 A40120	CAMP-responsive en
45	147	5.5	333	2 T26808	hypothetical prote

ALIGNMENTS

RESULT 1

A42140
box B-binding factor-2 - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C>Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 21-Jul-2000
C:Accession: A42140; S24542
R:Abel, T.; Bhatt, R.; Maniatis, T.
Genes Dev. 6, 466-480, 1992
A:Title: A Drosophila CREB/ATF transcriptional activator binds to both fat body- and liv
A:Reference number: A42140; MUID:92192458; PMID:1532159
A:Accession: A42140
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-515 <ABR>
A:Cross-references: EMBL:X64429; NID:g11063; PIDN:CAA45771.1; PID:g11064
A:Note: sequence extracted from NCB1 Backbone (NCBIP:88160)
C:Genetics:
A:Gene: FlyBase:Bbbf2
A:Cross-references: FlyBase:FBgn0004848
A:Superfamily: fos/jun DNA-binding domain homology
C:Keywords: DNA binding; nucleus; transcription regulation
F:432-475/Domain: fos/jun DNA-binding domain homology <FJD>

Query Match	16.8%	Score 445.5;	DB 2;	Length 515;
Best Local Similarity	33.8%	Pred. No. 8.4e-19;		
Matches 142;	Conservative 54;	Mismatches 117;	Indels 107;	Gaps 16;
QY 26 ETEALMYH---THFSELDEFSQNVLGQLLSDPFLSEKSEMEVEFSPSTPAPLQAEHS 82				
Db 111 DPKCLTFHVPPTHTATP-----ISRLSSNPALN-----TSVADLTRSSGL 149				
QY 83 YSL-SEERTQSPFTHAATSD-----SFNDEVESEKWL-----STEPS 122				
Db 150 OSLOAHQPHHGSGSHVVANLEHFQLPOLYDNCSSVSLRDGSMSPDICSDEIDE 209				
QY 123 ATIKKEPITEEPGLVPS-----VTITITATSP--FEKESPLDMNAGDSSCQ 171				
Db 210 SAIKDEPSPSSCAPTSQASSSQHQLSLNLAHQEMLEFEPKHCGLLTASSNSN-N 268				
QY 172 TLIPK-----TKLEPHEVDQFLNFPKASVDQLHLPTPPPS-HSSDSE 215				
Db 269 SLIKSQORHEQLGDNLMAKWEIKSEKQSTNSNGSHAGYGIPTTPPSLPSDDSE 328				
QY 216 GSLSPN-----PRL-----HPFSLSQASHVPR-----AMPGPSALST-- 248				
Db 329 GNLSPHLFRHCRPTQPPFLMPLIQFAVSHPYGSAATAITSSSGSASASGSSSTVT 388				
QY 249 -----SPLLTAPKLOSGPLVLTTEBKRTLVAEGYPIPTKPLTSEKALKKIREK 301				
Db 389 TTRQPIHTPLISSQPK-GSTGLTLLTEBKRTLVAEGYPIPTKPLTSEKALKKIREK 447				
QY 302 IKNKISAQESRRKKKEYMDSLEKVKESCSTENLELRKKVLENTNRTLLOQLQKLTIV 361				

Matches 94; Conservative 58; Mismatches 137; Indels 123; Gaps 18;

QY 1 MEVLSEGSQVLOWDRKLSLSEPEGETEALMYHTHPSSELLDFSONVLQQL-LSDPFLSE 59
Db 3 METVES-----QDRSV-----THSVAEHSLLHMQTQISVPTLAQSVAGSGTGR 48
QY 60 KSESVEVEPSPTSPA-----PLIAQHSYSLSSEPRQOS--PETHAATSDSFNDEEVESEK 113
Db 49 GSPATVLQVPSQTVQVGVQLTHP-SVIGSPQIQTVQVATIAETDSDASEVIDSHK 107
QY 114 W-----YLSTEPSATIKKEPITEEO--PPGL-----VP-SVTLTITTA---- 148
Db 108 RREILSRPSYRKILNLSDDVGPIKIEEKEEGTTPNIATWAVPISVIQTSTGQYI 167
QY 149 -----ISTPFKEESPLD-----MNAGGSSCQTLIPKIKLPEHVDQFLNFSPEA 195
Db 168 ATAOGGTIQISNPGSDGVQGLQALTMWNSGAPPPGATIVQYAAQSGADGTQOF--FVPGSQ 225
QY 196 SV---DQLHLPPTPPSSHSSDSEGLSPNRLHPSLSQAHSVPVAMPGRGPSALSTSPLL 252
Db 226 VVQDEETDLAP-----SHWAAATGDM-----PTYQIRAPTTALPGVVMASPGSL 272
QY 253 TAPHKLGQSLVLTTEEEKRTLVAEGYPIPTKPLTKSEKALKIRKIKNKISQAQESR 312
Db 273 HSPQQL-----ASEATRKEELRLMKNRAREAR 301
QY 313 RKKEYSMSLEKVVESCSTENLELRKKVEULENTNRTLLQLOKLOTLVMGK 364
Db 302 RKKEVYKCLE-----NRVAVLSSQNKLTIEELKALKLDLYCHK 339
RESULT 10
JC4028
activating transcription factor 2 - African clawed frog
N;Alternate names: cyclic AMP-response element-binding protein
C;Species: Xenopus laevis (African clawed frog)
C;Date: 13-Jun-1995 #sequence_revision 14-Jul-1995 #text_change 26-Aug-1999
C;Accession: JC4028
R;Villarreal, X.C.; Richter, J.D.
Gene 153, 225-229, 1995
A;Title: Analysis of ATF2 gene expression during early Xenopus laevis developme nt.
A;Reference number: JC4028; MUID:95180723; PMID:7875593
A;Accession: JC4028
A;Molecule type: mRNA
A;Residues: 1-486 <VIL>
A;Cross-references: GB:U16158; NID:988779; PIDN:AAA69518.1; PID:g710326
C;Comment: This protein is a sequence-specific DNA-binding protein that mediates transcr
Superfamily: cAMP response element-binding protein 1; fos/jun DNA-binding domain homol
Keywords: leucine zipper; phosphoprotein; transcription regulation
F:329-369/Region: leucine zipper motif
F:329-369/Domain: fos/jun DNA-binding domain homology <FJD>
F:332-364/Region: basic
F:82/Binding site: phosphate (Ser) (covalent) (by protein kinase A) #status predicted
F:102/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predicted

Query Match 6.1%; Score 162.5; DB 2; Length 486;
Best Local Similarity 21.5%; Pred. NO. 0.03;
Matches 101; Conservative 68; Mismatches 141; Indels 159; Gaps 24;

QY 52 LSDPF-----LSEKSEMEVEPSPTSPAPLIQAE-HSVLSLEERTQSPFTH--AATSD 102
Db 70 LASPPENEFKASEEDKKQPLDLSPLA-TPIIRSKIEFVSVVETTHQDSPLPHPESTTSD 128
QY 103 SFNDEVESEKWLSTEPSATIKKEPITEEQPGL-VPSVTLT-----ITAIPTPEK 155
Db 129 QHNQVPL-----AQAQPTSAIV-----PASLQVNPVLLTSSDSVLIQAVPSP 175
QY 156 RESPLDMNAGDSSCQTL-----IPKIKLEPHEVDQFLNFSPEASVDQLHLP--- 203
Db 176 SSTVITQAPSTNSSIVTLPGPFVLLHLPLSGQTMFVAI-----PASITSSNVHVA 228
QY 204 -----PTPSSSHSSDSEGLSPNRLHPSLSQAHSVPVAMPGRGPSALST 248

Db 229 PLLRLPLTWVPVPGIPGPSSQPQVQSEAKMRNA-----AWTQOHPV---TNGDTV--- 277
QY 249 SPLLPAPHKLGQSGPLVLTEEKR-----TLVAEGYPIPTK-LPLTKS----- 290
Db 278 -----KGHSQFAITQSESRPQSLQQPATSTTETRPSAQPTPOTQSTSGRRRAA 329
QY 291 -----BEKALKIRRKIKNKISAOESRRKKKYMDSLEKKVESCSTENLELRKKVEULENT 346
Db 330 NEDPDEKRSKIQR---NRAAASRCRQKRWQVQVQSELEKKADELISLNGYLQNEVTLRR- 385
QY 347 NRTLQLOKLOTLVMGKVSRCKLAGTGTGTCMLMVVLCPAVAFSGFPGQGYGPSATK 406
Db 386 -----EVAQLKQLL--AHKDCPTAMQKKS-----GHSADK 416
QY 407 -----MALPSQHPUSEPYTASVVRNLLIYEHAPELESSESSPASTGE 449
Db 417 DDSSEDISVP-----SSPTEAI-----QHSVSSTSGVSTSK 450
RESULT 11
TVQJUN
transforming protein jun - Japanese quail
A;Species: Coturnix coturnix japonica (Japanese quail)
C;Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 18-Jun-1999
C;Accession: S05963
R;Brun, G.; La Vista, N.; Dangi, J.P.; Castellazzi, M.
Nucleic Acids Res. 17, 6393, 1989
A;Title: Nucleotide sequence of the quail c-jun protooncogene.
A;Reference number: S05963; MUID:89366673; PMID:2505235
A;Accession: S05963
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-313 <BRU>
A;Cross-references: EMBL:X15547; NID:962639; PIDN:CAA33553.1; PID:g62640
C;Genetics:
A;Gene: jun
C;Superfamily: jun transforming protein; fos/jun DNA-binding domain homology
C;Keywords: DNA binding; leucine zipper; oncogene; transcription regulation;
F:229-269/Domain: fos/jun DNA-binding domain homology <FJD>
F:262-290/Region: leucine zipper motif

Query Match 6.1%; Score 162; DB 1; Length 313;
Best Local Similarity 22.5%; Pred. No. 0.018;
Matches 72; Conservative 66; Mismatches 128; Indels 54; Gaps 12;

QY 88 EPRTQSPFTHAATSDSFNDEEVESEKWLSTEPSATIKKEPITEEQPGLVPSVTLTIT 147
Db 11 EDALNASTAPPESAYGYNNAKVLKQSWTLNLSDFASSLK--PHLRKNADIITSPDVGLL 68
QY 148 AISTPFKEESPLDMNAG-----GDSSCQTLIPKIKLEPHE--VDQFLNFSPEASVDQLH 201
Db 69 KLASP-ELERLIQSSNGLITTTPTPTQFLCPKNVTDQEGFAEGFVR-----ALAEHL 121
QY 202 LPPTPPS--SHSSSEGLSP-----NPRLHPSLSQAHSVPVRA--MPRPSALS 247
Db 122 NQNTLPSVTSAAQPVSGMGPVSSMAGGSGFNTSLH-----SEPPVYANLSNFPNALN 175
QY 248 TSPLLTA-----PHKLOGSP-----LVLTEERKTLVAEGYPIPTKPLTKSEKAL 295
Db 176 SAPNVNANGMGAPOHHINPQMPVOHPRLOALKEEPQVPEMPGETPPLSPIDMESQRI 235
QY 296 KKIRKIKNKISAOESRRKKKYMDSLEKKVESCSTENLELRKKVEULENTNRTLLQLO 355
Db 236 KAERKMRNRIAAKCRKRKLRIARLEEKVTKLKAQNSE-----LASTANMLREQVA 288
QY 356 KLOTLVMGKVSRCKLAGTQ 375
Db 289 QLKQKVMNHNHNSGCOLMLTQ 308
RESULT 12
S36101
cAMP response element-binding protein epsilon - mouse

```

Db   229 TG-----EGQKKGRSP-----ADKESRLKLLRNRSVAQQARERKKAYLI 270
      :|          :|          :|          :|          :|          :|          :|
Qy   321 SLEKKVESCSTENLRKKVEVLNTNRTLLQOLQKLQTLVMGKVSRCTCKLAGTGTG 378
      ||:         |:         |:         |:         |:         |:         |:
Db   271 DLETRVKDLKKSELKERLSTLTQENOMLRQLK-----NTTASRRGSNNGT 318
      :|          :|          :|          :|          :|          :|          :|

RESULT 14
TVFVNJ
transforming protein jun - avian sarcoma virus 17
C:Species: avian sarcoma virus 17
C>Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 16-Feb-1999
C:Accession: A29870
R:Maki, Y.; Bos, T.J.; Davis, C.; Starbuck, M.; Vogt, P.K.
Proc. Natl. Acad. Sci. U.S.A. 84, 2848-2852, 1987
A>Title: Avian sarcoma virus 17 carries the jun oncogene.
A:Reference number: A29870; MUID:87204225; PMID:3033666
A:Accession: A29870
A:Molecule type: DNA
A:Residues: 1-296 <MAK>
C:Comment: This protein is translated as a gag-jun polypeptide.
C:Genetics:
A:Gene: jun
C:Superfamily: jun transforming protein; fos/jun DNA-binding domain homolog
C:Keywords: DNA binding; leucine zipper; oncogene; transcription regulation
F:212-252/Domain: fos/jun DNA-binding domain homology <FUD>
F:245-273/Region: leucine zipper motif

Query Match           6.0%; Score 159.5; DB 1; Length 296;
Best Local Similarity 23.0%; Pred. No. 0.023;
Matches 65; Conservative 45; Mismatches 87; Indels 85; Gaps

Qy   117 STEPPSATIKKEPITEBQ----PPGLVPDV-----TL-TITAIPTPEKEESPDMNAG 155
      :||:         |:         |:         |:         |:         |:         |:
Db   72 TTPTQTPLCPKNVTDEQGFAGFVRAALAEHLNQNTLPSTSAAPVSGMAPVSSMAG 131
      :|          :|          :|          :|          :|          :|          :|
Qy   166 GDSSCQTILPKILEPHEVDQFLNFSPEAS-----VDQLHLPPTPSHSDD 213
      :|          :|          :|          :|          :|          :|          :|
Db   132 GGSEFNSTL-----HSEPVPYANLSNFNALNSAPNYANNRMGYAPQHINPQPVQH--- 184
      |||         |:         |:         |:         |:         |:         |:
Qy   214 SEGSLPNRPLRPFLSLSQAHPVPMRPGPSALSPTSPLLTAHKLOGSGPLVLTEEBKRT 2737
      |||         |:         |:         |:         |:         |:         |:
Db   185 -----PRLQ--ALKXEPQTVEMFG-----ETPL-----200-----
      |||         |:         |:         |:         |:         |:         |:
Qy   274 LVAEGYPITKLPDTKBEBAKKIRKINKISQAQRKRKKKEYMDLSLEKKVESCSTEN 333
      |||         |:         |:         |:         |:         |:         |:
Db   208 -----FPIDMSQEIRKAERKMNRNRIAAKSRRKKLERIARLEEKVTKLKAQN 256
      |||         |:         |:         |:         |:         |:         |:
Qy   334 LELRKVVLENTNRTLLOQLQKLQTLVMGKVSRCTCKLAGTG 375
      :|          :|          :|          :|          :|          :|          :|
Db   257 SE-----LASTANMUREQVAQLQKVMNHVNSGCQLMLTG 291
      :|          :|          :|          :|          :|          :|          :|

RESULT 15
A37944
cAMP response element-binding protein alpha - mouse
C:Species: Mus musculus (house mouse)
C>Date: 12-Jul-1991 #sequence_revision 12-Jul-1991 #text_change 05-Nov-1991
C:Accession: A37944
R:Foulkes, N.S.; Borrelli, E.; Sassone-Corsi, P.
Cell 64, 739-749, 1991
A>Title: CREM gene: use of alternative DNA-binding domains generates mult
A:Reference number: A37944; MUID:91145594; PMID:1847666
A:Accession: A37944
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-229 <FOU>
A:Cross-references: GB:M60285; NID:g1927716; PIDN:AAAL7497.1; PID:g1927719
C:Superfamily: fos/jun DNA-binding domain homology
C:Keywords: alternative splicing; DNA binding; transcription regulation
F:163-206/Domain: fos/jun DNA-binding domain homology <FUD>
```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 5, 2003, 06:17:30 ; Search time 32 Seconds
(without alignments)
675.286 Million cell updates/sec

Title: US-09-884-566A-2

Perfect score: 2650

Sequence: 1 MEVLSEGSQSLQWDRKLSK.....LEGNETLKVVLERRVNVATF 521

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

11 number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	445.5	16.8	515	1 CRBA_DROME	P29747 drosophila
2	252	9.5	699	2 ATFB_MOUSE	O35451 mus musculus
3	223.5	8.4	703	1 AT6B_HUMAN	Q95941 h cyclic-am
4	212	8.0	670	1 AT6A_HUMAN	P18850 homo sapien
5	169	6.4	314	1 APL_SERCA	P54864 serinus can
6	167.5	6.3	741	1 NFL1_MOUSE	Q61985 mus musculus
7	162.5	6.1	341	1 CREM_RAT	Q03061 rattus norv
8	162	6.1	313	1 APL_COTJA	P12981 coturnix co
9	159.5	6.0	296	1 TUUN_AVIS1	P05411 avian sarco
10	158.5	6.0	271	1 ATF1_HUMAN	P18846 homo sapien
11	158	6.0	344	1 CREM_CANFA	P79145 canis famil
12	157.5	5.9	269	1 ATF1_MOUSE	P81269 mus musculus
13	157	5.9	772	1 NFL1_HUMAN	Q14494 h nuclear f
14	156.5	5.9	310	1 APL_CHICK	P18870 gallus gall
15	152.5	5.8	373	1 NFE2_HUMAN	Q18621 homo sapien
16	152	5.7	308	1 JUNB_CYPCA	P19703 cyprinus ca
17	152	5.7	334	1 APL_MOUSE	P05627 mus musculus
18	152	5.7	489	1 YEN5_YEAST	P39970 saccharomyc
19	150.5	5.7	331	1 APL_FIG	P56432 sus scrofa
20	150	5.7	341	1 JUND_MOUSE	P15066 sus musculus
21	150	5.7	341	1 JUND_RAT	P52909 rattus norv
22	150	5.7	605	1 NFL2_HUMAN	Q16236 homo sapien
23	150	5.7	679	1 YIS3_YEAST	P40563 saccharomyc
24	149.5	5.6	325	1 CREB_BOVIN	P27925 bos taurus
25	149.5	5.6	1844	1 POLR_TYMYA	P20128 turnip yell
26	149	5.6	334	1 APL_RAT	P17325 rattus norv
27	149	5.6	494	1 ATF7_HUMAN	P17544 homo sapien
28	148.5	5.6	487	1 ATF2_RAT	Q00969 rattus norv
29	147.5	5.6	341	1 CREM_MOUSE	P27699 mus musculus
30	147.5	5.6	487	1 ATF2_HUMAN	P15336 homo sapien
31	147.5	5.6	487	1 ATF2_MOUSE	P16951 mus musculus
32	146.5	5.5	331	1 APL_HUMAN	P05412 homo sapien
33	146.5	5.5	743	1 TPE3_HUMAN	P19532 homo sapien

Query Match 16.8%; Score 445.5; DB 1; Length 515;

ALIGNMENTS

RESULT 1

CRBA_DROME
ID CRBA_DROME STANDARD; PRT; 515 AA.
AC P29747;

DT 01-APR-1993 (Rel. 25, Created)

DT 01-APR-1993 (Rel. 25, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Cyclic-AMP response element binding protein A (Box B binding factor-2)

DE (BBF-2).

GN CREBA OR BBF2.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;

OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;

OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Oregon-R;

RX MEDLINE=92192458; PubMed=1532159;

RA Abel T., Bhatt R., Maniatis T.;

RT "A Drosophila CREB/ATF transcriptional activator binds to both fat

body- and liver-specific regulatory elements.";

RL Genes Dev. 6:466-480(1992).

CC -!- FUNCTION: TRANSCRIPTIONAL ACTIVATOR. BINDS TO FAT BODY-SPECIFIC

ENHANCERS OF ALCOHOL DEHYDROGENASE (ADH) AND YOLK PROTEIN GENES.

CC BBF-2 MAY PLAY A ROLE IN FAT BODY GENE EXPRESSION. IT BINDS THE

CONSENSUS SEQUENCE 5'-T(A/C)NACGTAN(T/G)C-3'.

CC -!- SUBUNIT: MAY BIND DNA AS HETERODIMERS WITH OTHER BZIP PROTEINS.

CC -!- SUBCELLULAR LOCATION: Nuclear.

CC -!- TISSUE SPECIFICITY: IN ALL CELL TYPES EXAMINED.

CC -!- DEVELOPMENTAL STAGE: PRESENT THROUGHOUT DEVELOPMENT.

CC -!- SIMILARITY: BELONGS TO THE BZIP FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; X64429; CAA45771.1; -.

PIR; S24542; S24542.

PIR; A42140; A42140.

TRANSFAC; T01603; -.

FlyBase; FBgn0004396; Creba.

InterPro; IPR004827; TF_BZIP.

Pfam; PF00170; BZIP; 1.

SMART; SM00338; BRLZ; 1.

PROSITE; PS00036; BZIP_BASIC; 1.

Transcription regulation; Activator; DNA-binding; Nuclear protein.

DNA BIND 442 462 BASIC MOTIF.

DOMAIN 468 503 LEUCINE-ZIPPER.

SQ SEQUENCE 515 AA; 56528 MW; 0808FB9655200223 CRC64;

FT	CARBOHYD	676	676	N-LINKED (GLCNAC...) (POTENTIAL).
FT	VARSPLIC	28	31	GLQN -> D (IN ISOFORM 1).
FT	CONFLICT	3	3	E -> D (IN REF. 2).
FT	CONFLICT	329	330	QQ -> HE (IN REF. 2).
FT	CONFLICT	520	520	V -> D (IN REF. 4; AAC50888).
FT	CONFLICT	600	600	D -> G (IN REF. 1).
FT	CONFLICT	600	600	D -> G (IN REF. 1).
SQ	SEQUENCE	703 AA;	76709 MW;	B04C7B2337D83F82 CRC64;
Query Match 8.4%; Score 223.5; DB 1; Length 703;				
Best Local Similarity 23.5%; Pred. No. 2.2e-05;				
Matches 135; Conservative 78; Mismatches 216; Indels 145; Gaps 24;				
Qy	18	LSEUSEP-----GETEALMTHFSELLDFEQNVQLQLSDPFL-----SEKS 61		
Db	7	LSEADPTFRFTDNLSPEDWGLQNSTLYSG-LDEVAEEQT-QLFRCPEDQVPPFGSSLD 64		
Qy	62	ESMEVEPPTSPAPLIQAEHSYLSLEERTQSPFTHTAATSDSFNDEEVESEKWLSTTFPP 121		
Db	65	VQMDVSPS-EPWELLPDPDLOVKSEP--SSPCSSSLU-----SESRLLSTEPS 112		
Qy	122	SATIKKEPTEQPGLVPSVTLTITATSTPEKESPLDMNAGDSSQTLIPKIKLEP 181		
Db	113	SEALGVGEVLHVKTSLAPPLCLGDDPTSSFTQINVIPTSDSDSVQT-----KTEP 167		
Qy	182	HEVDQFLNFPKESVDQ-----LHLPPTP-----P 207		
Db	168	VSPCSSVNSEASLLSADSSQAFIGEVLVETESLSPSGCLLWDVPAPSLGAVQISMP 227		
Qy	208	SHSSDSSEGLSPNRLH--PFLSQAHSPVRAMRG-----PSALSTSPLLTAP 255		
Db	228	SLDGSGLKALPRKPLQPKVLLTVPMPSRAVPSPTVLLQSLVQPPVPFVLIQQA 287		
Qy	256	HKLQSGPL-VLTHEKRTLVAGVPPIPTKLPTKSEKALKIRKIRKIKKISAOESRRK 314		
Db	288	IRVQEGPAPSLPRPERKIVPA--PMPGNSCPPEVDKLLKXQQRMIKNRESACQSRK 345		
Qy	315	KKEYMDSLEKKVESCSTENLEIRKKVEVLNENTRLTLOQLKLTVMGKVERTCKLAGT 374		
Db	346	KKEYLQGLARQVLADNQQLRRENAALRLLEALLAENSELK--LG-----SGN 394		
Qy	375	QTGTCLMVVLCFAVAFSGFFQGYGPYPATKWLPSQHPLS-----PYPTASVVRSN 428		
Db	395	RKVVICMVFLLFTAFNFG-----PVSISEPSPAPISPRMKGEPQP-----RRH 438		
Qy	429	LLIYEHAPLESSESPASTGELGMDRGSSLLRASSGLEALPE---VDLPFLFLISNETSL 485		
Db	439	LLGFSQEPVQ-----GVEPLQSGQKPEQPSPTDQPSF--SNLTAF 480		
Qy	486	-----EKSVLLEQLHLVSS---KLEGNETLKVVE 512		
Db	481	PGGAKELLRLDLQFLSLDCRHFNRFTESURLAD 514		
RESULT 4				
AT6A HUMAN STANDARD; PRT; 670 AA.				
AC	P18850; O15139; Q9UEC9;			
DT	01-NOV-1990 (Rel. 16, Created)			
DT	30-MAY-2000 (Rel. 39, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Cyclic-AMP-dependent transcription factor ATF-6 alpha (Activating			
DE	transcription factor 6 alpha) (ATF6-alpha).			
GN	ATF6.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OC	NCBI_Taxid=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Cervical carcinoma;			
RX	MEDLINE=97415575; PubMed=9271374;			
RA	Zhu C., Johansen F.E., Prywes R.;			
RT	"Interaction of ATF6 and serum response factor."			

RL	Mol. Cell. Biol. 17:4957-4965(1997).
RN	[2]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Cervical carcinoma;
RX	MEDLINE=99057946; PubMed=9837962;
RA	Yoshida H., Haze K., Yanagi H., Yura T., Mori K.;
RT	"Identification of the cis-acting endoplasmic reticulum stress
RT	response element responsible for transcriptional induction of
RT	mammalian glucose-regulated proteins; involvement of basic-leucine
RT	zipper transcription factors."
RL	J. Biol. Chem. 273:33741-33749(1998).
RN	[3]
RP	SEQUENCE OF 302-369 FROM N.A.
RX	MEDLINE=90195187; PubMed=2516827;
RA	Hai T., Liu F., Coukos W.J., Green M.R.;
RT	"Transcription factor ATF CDNA clones: an extensive family of leucine
RT	zipper proteins able to selectively form DNA-binding heterodimers."
RL	Genes Dev. 3:2083-2090(1989).
RN	[4]
RP	ERRATUM.
RA	Hai T., Liu F., Coukos W.J., Green M.R.;
RL	Genes Dev. 4:682-682(1990).
RN	[5]
RP	SEQUENCE OF 1-202 FROM N.A.
RC	TISSUE=Pancreas;
RA	Strausberg R.;
RL	Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN	[6]
RP	CHARACTERIZATION.
RX	MEDLINE=20032174; PubMed=10564271;
RA	Haze K., Yoshida H., Yanagi H., Yura T., Mori K.;
RT	"Mammalian transcription factor ATF6 is synthesized as a transmembrane
RT	protein and activated by proteolysis in response to endoplasmic
RT	reticulum stress."
RL	Mol. Biol. Cell. 10:3787-3799(1999).
RN	[7]
RP	PROCESSING BY PS1 AND PS2, AND MUTAGENESIS OF ARG-415; ARG-416;
RP	LEU-419; ASN-391 AND PRO-394.
RX	MEDLINE=2111045; PubMed=11163209;
RA	Ye J., Rawson R.B., Komuro R., Chen X., Dave U.P., Prywes R.,
RA	Brown M.S., Goldstein J.L.;
RT	"ER stress induces cleavage of membrane-bound ATF6 by the same
RT	proteases that process SREBPs."
RL	Mol. Cell. 6:1355-1364(2000).
RN	[8]
RP	FUNCTION.
RX	MEDLINE=21106370; PubMed=11158310;
RA	Yoshida H., Okada T., Haze K., Yanagi H., Yura T., Negishi M.,
RA	Mori K.;
RT	"Endoplasmic reticulum stress-induced formation of transcription
RT	factor complex ERSP including NF-Y (CBF) and activating transcription
RT	factors 6alpha and 6beta that activates the mammalian unfolded
RT	protein response."
RL	Mol. Cell. Biol. 21:1239-1248(2001).
RN	[9]
RP	REVIEW.
RX	MEDLINE=21376119; PubMed=11483355;
RA	Hai T., Hartman M.G.;
RT	"The molecular biology and nomenclature of the activating
RT	transcription factor/CAMP responsive element binding family of
RT	transcription factors: activating transcription factor proteins and
RT	homeostasis."
RL	Gene 273:1-11(2001).
RN	CC
CC	-!- FUNCTION: TRANSCRIPTION FACTOR THAT ACTS DURING ENDOPLASMIC
CC	RETICULUM STRESS BY ACTIVATING UNFOLDED PROTEIN RESPONSE TARGET
CC	GENES. BINDS DNA ON THE 5'-CCAC[GA]-3' HALF OF THE ER STRESS
CC	RESPONSE ELEMENT (ERSE) (5'-CCAAT-N9-CCAC[GA]-3') AND OF ERSE II
CC	(5'-ATTGG-N-CCACG-3'). BINDING TO ERSE REQUIRES BINDING OF NF-Y TO
CC	ERSE. COULD ALSO BE INVOLVED IN ACTIVATION OF TRANSCRIPTION BY THE
CC	SERUM RESPONSE FACTOR.
CC	-!- SUBUNIT: HOMODIMER AND HETERODIMER WITH ATF6-BETA. THE DIMER
CC	INTERACTS WITH THE NUCLEAR TRANSCRIPTION FACTOR Y (NF-Y) TRIMER
CC	THROUGH DIRECT BINDING TO NF-Y SUBUNIT C (NF-YC). INTERACTS ALSO

Db 503 DUEAEGAVGYPEYKSCRMYSQPSQLSCPLYLEHVHCHNHTYNWASALDSADLPSPS 562
Qy 265 VLTEEEKRTL-----AEGYIP-----TKLPLTKSEK 293
Db 563 TLKGSKEKQADFLDKMSRDEHRAMKIPPTNDKIINLPVEEFNELLKSKYLSEAQLS 622
Qy 294 ALKKTRRKTKKISAEQSRKKKEYMDSLEKKEVSCSTENLE-LKKKVEVLENTNRLTLO 352
Db 623 LIRDIRRRGNKNAQNCRRKLDLTILNLERDVEDLQDKARLLREKVEFL-----RSURQ 678
Qy 353 QLOKQLTVMGKVSRTCKLAGTGTCTCLMVVLCFAVAFGSPFQGVG-PYPSATKMALPS 411
Db 679 MKQKVQSL-----YQEVGRRLDERGPRYS-----PS 705
Qy 412 QHPLSEPYTASV 424
Db 706 QVALQVAGDGSVL 718
RESULT 7
ID CREAM RAT
AC Q03061; STANDARD; PRT; 341 AA.
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE CAMP-responsive element modulator.
GN CREAM.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=93096608; PubMed=1461747;
RA Meyer T.E., Habener J.F.;
RT "Cyclic AMP response element binding protein CREB and modulator
protein CREM are products of distinct genes.";
RL Nucleic Acids Res. 20:6106-6106(1992).
CC -!- FUNCTION: THIS PROTEIN BINDS THE CAMP RESPONSE ELEMENT (CRE).
CC ALPHA, BETA, AND GAMMA ISOFORMS ARE ANTAGONISTS OF THE CAMP
CC TRANSCRIPTIONAL RESPONSE, WHILE THE DELTA ISOFORM IS AN ACTIVATOR.
CC -!- SUBUNIT: BINDS DNA AS A DIMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- ALTERNATIVE PRODUCTS: 4 ISOFORMS: ALPHA, BETA, GAMMA AND DELTA
(SHOWN HERE); ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -!- PTM: STIMULATED BY PHOSPHORYLATION (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE BZIP FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z15158; CA78857.1; -.
CC PIR; S26686; S26686.
CC HSSP; P03069; 2DGC.
CC
CC TRANSFAC; T01919; -.
CC InterPro; IPR001630; Leuzip CREB.
CC InterPro; IPR004827; TF_BZIP.
CC InterPro; IPR003102; pKID.
CC Pfam; PF00170; bzip; 1.
CC Pfam; PF02173; pKID; 1.
CC PRINTS; PR00041; LEUZIPRCREB.
CC SMART; SM00338; BRLZ; 1.
CC PROSITE; PS00036; BZIP_BASIC; 1.
KW Transcription regulation; DNA-binding; Activator; Repressor;

KW Phosphorylation; Nuclear protein; Alternative splicing.
FT DNA_BIND 284 305 BASIC MOTIF (BY SIMILARITY).
FT DOMAIN 311 332 LEUCINE-ZIPPER (BY SIMILARITY).
SQ SEQUENCE 341 AA; 36628 MW; 1DE752148670125E CRC64;
Query Match 6.1%; Score 162.5; DB 1; Length 341;
Best Local Similarity 22.8%; Pred. No. 0.02;
Matches 94; Conservative 58; Mismatches 137; Indels 123; Gaps 18;
Qy 1 MEVLFGSQSVQLMDRKULSELSEPGETEALMYHFSSELLDEFQNVLGQL-LSDPFLSE 59
Db 3 METVES-----QDRSV-----THSVAEHSLSLHMGTGQISVPTLQAVSVAGSGTGR 48
Qy 60 KSESENEVPSPTSPA-----PLIQAEHSYSLSEEPQTOS--PETHAATSDSFNDEVESEK 113
Db 49 GSPAVTLVQLPSGQTVQGVQVITQTPH-SVQSPQIQTVQVATIAETDDSDASEVIDSHK 107
Qy 114 W-----YLSTFFPSATIKKEPITERQ--PPGL-----VP-SVLTITTA--- 148
Db 108 RREILSRPSYRKILNELSSDVPGIPKIEEKESEEGTPPNATWAVPTSIYQTSTGOYI 167
Qy 149 -----ISTPFKEBSPLD-----MNAGDSSCQTLPKIKLEPHEVDQFLNFSPEKA 195
Db 168 AIAQGGTIQISNPGSDGVQGLQALTMTNMGAPPGATIVQYAAQADGTQGF--FVPGSQ 225
Qy 196 SV---DQLHLPTPPSSHSSDSEGLSNPRLHPFSLSQAHSPVPMRPGPSALSTSPLL 252
Db 226 VVQDEETDLAP-----SHMAATGDM-----PTYQIRAPTTALPQGVVMAASPGSL 272
Qy 253 TAPHKQSGPLVLTEBEKRTLVAEGYPIPTKLPKTKSEKALKKIRKIKNKISQAQSR 312
Db 273 HSPQQL-----AEEATRKRELMLMKREARECR 301
Qy 313 RKKKEYMDSLEKKVSCSTENLEKKEVLENTNRLTLOKQLQTLVMGK 364
Db 302 RKKKEYVKCLE-----NRVAVLESQNKLTIELKALKDLYCHK 339
RESULT 8
ID API_COTJA STANDARD; PRT; 313 AA.
AC P12981;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Transcription factor AP-1 (Proto-oncogene c-jun).
GN JUN.
OS Coturnix coturnix japonica (Japanese quail).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Coturnix.
OX NCBI_TaxID=93934;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89366673; PubMed=2505235;
RA Brun G., la Vista N., Dangy J.-P., Castellazzi M.;
RT "Nucleotide sequence of the quail c-jun protooncogene.";
RL Nucleic Acids Res. 17:6393-6393(1989).
CC -!- FUNCTION: TRANSCRIPTION FACTOR THAT BINDS AND RECOGNIZE THE
CC ENHANCER DNA SEQUENCE: TGA(C/G)TCA
CC -!- SUBUNIT: INTERACTS WITH C-FOS TO FORM A DIMER.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: BELONGS TO THE BZIP FAMILY. JUN SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X15547; CAA33553.1; -.


```

Db 251 ----PTYQIRAPTALPGQVMAASPGSLHSPQOL----- 281
Qy 284 KLPLTKSEKALKIRRKIKNKISQAQSRKKKKEYMDSLEKKVESCSTENLELRKKVEVL 343
Db 282 -----AEBATRKRLRLMKNEARECRKKKEYVKCLE-----NRVAVL 321
Qy 344 EVNTNRTLQOLQKLOLTVMGK 364
Db 322 ENQNTLIEELKALKDLYCHK 342

RESULT 12
ATF1_MOUSE
ID ATF1_MOUSE STANDARD; PRT; 269 AA.
AC P81369;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Cyclic-AMP-dependent transcription factor ATF-1 (Activating
transcription factor 1) (TCF-ATF1).
ATF1.
Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=92176656; PubMed=1531847;
RA Lei M.-R.; Chung C.-S.; Liou M.-L.; Wu M.; Li W.-F.; Hsueh Y.-P.;
RA Lai M.-Z.;
RT "Isolation and characterization of nuclear proteins that bind to T
cell receptor V beta decamer motif."
RL J. Immunol. 148:1906-1912(1992).
CC -!- FUNCTION: THIS PROTEIN BINDS THE CAMP RESPONSE ELEMENT (CRE)
(CONSENSUS: 5'GTGACGT(A/C)(A/G)-3'), A SEQUENCE PRESENT IN MANY
VIRAL AND CELLULAR PROMOTERS. BINDS TO THE TAX-RESPONSIVE ELEMENT
(TRE) OF HTLV-I.
CC -!- SUBUNIT: BINDS DNA AS A DIMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: BELONGS TO THE BZIP FAMILY. ATF SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; M63725; AAA40395.1; -.
CC MGD; MGI:1298366; Atf1.
CC InterPro; IPR001630; Leuzip_CREB.
CC InterPro; IPR004827; TF_BZIP.
CC InterPro; IPR003102; pKID.
CC Pfam; PF00170; bzip; 1.
CC Pfam; PF02173; pKID; 1.
CC PRINTS; PR00041; LEUZIPPCREB.
CC SMART; SM00338; BRLZ; 1.
CC PROSITE; PS00036; BZIP_BASIC; 1.
KW Transcription regulation; DNA-binding; Activator; Nuclear protein.
FT DNA_BIND 214 233 BASIC MOTIF.
FT DOMAIN 239 260 LEUCINE-ZIPPER.
SQ SEQUENCE 269 AA; 29237 MW; 9885265159D64A0C CRC64;

Query Match 5.9%; Score 157.5; DB 1; Length 269;
Best Local Similarity 24.7%; Pred. No. 0.028;
Matches 81; Conservative 47; Mismatches 115; Indels 85; Gaps 16;

Qy 58 QSKSSENEV--EFSPTSPAPLI-QAEHSY-SLSREPTQSPFFTHAATSDSFNDEEVESEK 113
Db 4 SHKSNTTETASQPGSTVAGPHVSVIHQVSLSESESSQD-----SSDSIG-----SSQK 53
Qy 114 WY-LSHFPSS-----ATIKKEPITEEPPGLVPSVTLITAITSPFPKESPLDMNAGGD 167

```

```

Db 54 AHGILARRPSYRKILKOLSSDTRCKGEGENPSIS-AITSMSP-----ADIYQTSSQ 107
Qy 168 SSCQTLIPKIKLEPHEVQDFNFSPEASVDOLH-LPTTPSS-----HSDSEG 216
Db 108 -----YIAIAPNGALQLA--SPSTDGVQALQTLTWTNSSSTQOQTILQVAQTSQGO 157
Qy 217 SLSPNRLHPFSLSAHSPVRAMPRGPSALSTPLTAPHKLQSGPLVLTTEEKRTLVA 276
Db 158 ILVPSNQV-----VVQTASGDMQTYQIRTPSATSLPQTVMVTS----- 196
Qy 277 EGYPIPTKLPLTKSEKALKIRRKIKNKISQAQSRKKKKEYMDSLEKKVESCSTENLEL 336
Db 197 ---PVTLASQTTKTDPPQLRREIRLMKNREARECRKKKEYVKCLE----- 240
Qy 337 RKKEVLENTNRTLQOLQKLOLTVMGK 364
Db 241 -NRVAVLENQNTLIEELKALKDLYCHK 267

RESULT 13
NFLI_HUMAN
ID NFLI_HUMAN STANDARD; PRT; 772 AA.
AC Q14494; Q12877;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Nuclear factor erythroid 2 related factor 1 (NF-E2 related factor 1)
(NFE2-related factor 1) (Nuclear factor, erythroid derived 2, like 1)
(DE Transcription factor 1) (Transcription factor HBZ17) (Transcription
factor LCR-F1) (Locus control region-factor 1).
GN NFE2L1 OR NRF1 OR TCF11 OR HBZ17.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=95095252; PubMed=8001966;
RA Luna L.; Johnsen O.; Skartlien A.H.; Pedetour F.; Turc-Carel C.;
RA Prydz H.; Kolatse A.-B.;
RT "Molecular cloning of a putative novel human bzip transcription
factor on chromosome 17q22."
RL Genomics 22:553-562(1994).
RN [2]
SEQUENCE OF 326-772 FROM N.A.
RX MEDLINE=94310069; PubMed=8036168;
RA Caterina J.J.; Donze D.; Sun C.W.; Ciavatta D.J.; Townes T.M.;
RA "Cloning and functional characterization of LCR-F1; a bzip
transcription factor that activates erythroid-specific, human globin
gene expression."
RL Nucleic Acids Res. 22:2383-2391(1994).
CC -!- FUNCTION: ACTIVATES ERYTHROID-SPECIFIC, GLOBIN GENE EXPRESSION.
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- SIMILARITY: BELONGS TO THE BZIP FAMILY. CNC SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; X77366; CAA54555.1; -.
CC DR EMBL; U08853; AAA20466.1; -.
CC DR HSP; P34707; LSKN.
CC DR Genew; HGNC:7781; NFE2L1.
CC DR MIM; 163260; -.
CC DR InterPro; IPR004827; TF_BZIP.
CC DR SMART; SM00338; BRLZ; 1.
CC DR PROSITE; PS00036; BZIP_BASIC; 1.
KW Transcription regulation; Activator; DNA-binding; Nuclear protein.

```

```
FT DOMAIN 125 288 ASP/GLU-RICH (ACIDIC).
FT DOMAIN 496 517 POLY-SER.
FT DNA_BIND 659 674 BASIC MOTIF.
FT DOMAIN 682 704 LEUCINE-ZIPPER.
SQ SEQUENCE 772 AA; 84703 MW; C868807C6046BEFS CRC64;

Query Match
Best Local Similarity 5.9%; Score 157; DB 1; Length 772;
Matches 80; Conservative 62; Mismatches 135; Indels 66; Gaps 13;

Qy 28 EALMHTHPSSEL-LDFSONVLQGLSDPLSEKESMEVEPSPTSPAPLQAEHSYSL 86
Db 444 EALMDEISLMDLAIEGFPVQASQLEEFDSGLSD---SSHSPSSLSSESSSSS 500
Qy 87 EBPRTQSPTHAATSDSFNDEVESEKWLSTFPSTATIKKPIETEEQPGIPLVPSVTLTI 146
Db 501 SSSSSSSSSASSSSSSSEGAAGV---YSS-----DSETLDEEAGAV----- 542
Qy 147 TAISTPFKEEPLDMAGGSSCOTLIPKIKLEPHEVDQFLNFPKPEASVDOLHLPPTP 206
Db 543 -----GYQPEYSKFCRMSYQDPAQLSCLPYLEHVGH--NHTYNMAP--SALDSADLPPP- 592
Qy 207 PSSHSDSEGLSPNRLPHFPLSQAHSPVRAMPGRPSALSTSPILTAPHKLQSGPLVL 266
Db 593 ---SALKKGSKEKQADFLDKQMSRDEHRAAM---KIPFTNDKIINLP----- 634
Qy 267 TEEKRTLVAGYPIPTKPLTKSEKALKIRKIKNKISAEQSRKKKEYWDSLEKKV 326
Db 635 -----VEEFNELLSKYQLSEALSLTIRRRGKNAQAQNCRRKRLDTILNLERDV 686
Qy 327 ESCSTENLE-LRKKEVLENTNRTLLQQLQKLTLL---VMGKV 365
Db 687 EDLQDKARLLREKVEFL-----RSLRQMKQKQVLSYQEVFGRL 725

RESULT 14
APL_CHICK STANDARD; PRT; 310 AA.
AC APL18670;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Transcription factor AP-1 (Proto-oncogene c-jun).
GN JUN.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
NCBI_TaxID=9031;
[1]_TaxID=9031;
RP SEQUENCE FROM N.A.
RC TISSUE=Fibroblast;
RX MEDLINE=94167107; PubMed=2577867;
RA Nishimura T., Vogt P.K.;
RT "The avian cellular homolog of the oncogene jun.";
RL Oncogene 3:659-663(1988).
CC -!- FUNCTION: TRANSCRIPTION FACTOR THAT BINDS AND RECOGNIZE THE
CC ENHANCER DNA SEQUENCE: TGA(C/G)TCA.
CC -!- SUBUNIT: INTERACTS WITH C-FOS TO FORM A DIMER.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: BELONGS TO THE BZIP FAMILY. JUN SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; M57467; AAA48927.1; -.
DR HSSP; P05412; 1FOS.
DR TRANSFAC; T00134; -.
```

```
DR InterPro; IPR002112; Leuzip Jun.
DR InterPro; IPR004827; TF_bZIP.
DR Pfam; PF00170; BZIP; 1.
DR PRINTS; PR00043; LEUZIPPJUN.
DR SMART; SM00338; BRIZ; 1.
DR PROSITE; PS00036; BZIP_BASIC; 1.
KW Proto-oncogene; Transcription regulation; DNA-binding;
KW Nuclear protein.
FT DNA_BIND 236 255 BASIC MOTIF.
FT DOMAIN 259 287 LEUCINE-ZIPPER.
SQ SEQUENCE 310 AA; 33940 MW; 3EB3GB31F96C994E CRC64;

Query Match
Best Local Similarity 5.9%; Score 156.5; DB 1; Length 310;
Matches 72; Conservative 56; Mismatches 112; Indels 101; Gaps 12;

Qy 74 APLQAEHSYSLSEPRTPQSPFTHAATSDSFNDEEV-----ESEKWL-----S 117
Db 27 AKVLQSQMTLNSDAASLSKPHLRNKNADILTSVDGLLKLASPELERLIQSSNGLITT 86
Qy 118 TEFPSTATIKKPIETEEQ---PPGLVPSV-----TL-TITAISTPFKEESPLDMNAGG 166
Db 87 TPTPTQFLCPKNVTDEQGFAGFVRALAEHLNQNTLPSTSAAPQVSGMAPVSSMAGG 146
Qy 167 DSSCOTLIPKIKLEPHEVDQFLNFPKPEAS-----VDOLHLPPTPPSSHSSDS 214
Db 147 GSFNTSL-----HSEPPVYANLSNFNALSAPNANGMYPQHHINPQVQH----- 198
Qy 215 EGSLSNPRLPHFPLSQAHSPVRAMPGRPSALSTSPILTAPHKLQSGPLVLTEEEKRTL 274
Db 199 -----PRLQ--ALKERPQVPEMPG-----ETPPL----- 221
Qy 275 VAGYPIPTKPLTKSEKALKIRKIKNKISAEQSRKKKEYWDSLEKKKEYSCSTENL 334
Db 222 -----SPIDMESQERIKAERKMRNRIASCKRKLRIARLEEKVKTLKAQNS 271
Qy 335 ELRKKVEVLENTNRTLLQQLQKLTLLVMGKVSRTCKLAGTQ 375
Db 272 E-----LASTANMLREQVQLKOKVMNVHNSGCQLMLTQ 305

RESULT 15
NFE2_HUMAN STANDARD; PRT; 373 AA.
AC Q16621; Q07720;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Transcription factor NF-E2 45 kDa subunit (Nuclear factor, erythroid-
DE derived 2 45 kDa subunit) (P45 NF-E2) (Leucine zipper protein NF-E2).
GN NFE2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94068604; PubMed=8248255;
RA Chan J.Y., Han X.L., Kan Y.W.;
RT "Isolation of cDNA encoding the human NF-E2 protein.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:11366-11370(1993).
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93360994; PubMed=8355703;
RA Ney P.A., Andrews N.C., Jane S.M., Safer B., Purucker M.E.,
RA Weremowicz S., Goff S.C., Orkin S.H., Neinhuis A.W.;
RT "Purification of the human NF-E2 complex: cDNA cloning of the
RT hematopoietic cell-specific subunit and evidence for an associated
RT partner.";
RL Mol. Cell. Biol. 13:5604-5612(1993).
[3]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal liver;
```

Search completed: March 5, 2003, 06:19:23
Job time : 46 secs

```

RX MEDLINE=95241531; PubMed=7724591;
RA Pischetta C., Cocco S., Melis A., Marini M.G., Kan Y.W., Cao A.,
RT "Isolation of a differentially regulated splicing isoform of human
NF-E2.";
RL Proc. Natl. Acad. Sci. U.S.A. 92:3511-3515(1995).
CC -|- FUNCTION: REQUIRED FOR ACTIVITY AT THE LOCUS CONTROL REGION (LCR)
CC UPSTREAM OF THE GLOBIN GENE COMPLEXES. REQUIRES P18 NF-E2 FOR
CC BINDING TO THE NF-E2 MOTIF. MAY PLAY A ROLE IN ALL ASPECTS OF
CC HEMOGLOBIN PRODUCTION: GLOBIN SYNTHESIS, HEME SYNTHESIS, AND THE
CC PROCUREMENT OF IRON.
CC -|- SUBUNIT: HETERODIMER OF P45 NF-E2 AND P18 NF-E2.
CC -|- SUBCELLULAR LOCATION: Nuclear.
CC -|- TISSUE SPECIFICITY: EXPRESSED IN HEMATOPOIETIC CELLS AND ALSO IN
CC COLON AND TESTIS.
CC -|- SIMILARITY: BELONGS TO THE BZIP FAMILY. CNC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; L24122; AAA16118.1; -
DR EMBL; L13974; AAA35612.1; -
DR EMBL; S77763; AAB34115.1; -
DR HSSP; P34707; 1SKN.
DR TRANSFAC; T01440; -
DR Genew; HGNC:7780; NFE2.
DR MIM; 601490; -
DR InterPro; IPR004827; TF_BZIP.
DR Pfam; PF00170; BZIP; 1.
DR SMART; SM00338; BRLZ; 1.
DR PROSITE; PS00036; BZIP_BASIC; 1.
KW Transcription regulation; Activator; DNA-binding; Nuclear protein.
FT DOMAIN 59 62 POLY-PRO.
FT DOMAIN 77 82 POLY-PRO.
FT DNA_BIND 268 287 BASIC MOTIF.
FT DOMAIN 294 339 LEUCINE-ZIPPER.
FT CONFLICT 114 114 G -> A (IN REF. 2).
FT CONFLICT 214 214 A -> R (IN REF. 2).
FT CONFLICT 334 335 FO -> LE (IN REF. 2).
FT CONFLICT 334 335
SQ SEQUENCE 373 AA; 41472 MW; A9821170FB2ED67C CRC64;

Query Match 5.8%; Score 152.5; DB 1; Length 373;
Best Local Similarity 24.1%; Pred. NO. 0.08;
Matches 78; Conservative 30; Mismatches 104; Indels 111; Gaps 12;

QY 118 TFPFSATIKKEPTTEQPP-----GLVPSVTLTITAI-----STPPEKESPLDMNA 164
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 35 TELQGLNAPSESPFQAPAPYLGPPPTTYCPCSIHPDSGFLPPPPYSLPASTSHVPD 94
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 165 GGDSSCQTLIPKLEPHEVDQFLNFPKESVDQLHL-----PPTPPSSHSDS-- 214
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 95 PPYSYGNMAIPVSK--PLSLGSLIS-----EPLQDPLALLDIGLPAGPPKQEDPESDGL 148
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 215 -----EG-----SLSPNPRLH-----PFSLSQA 232
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 149 SLNYSDAESLELEGTEAGRRRSEVEMYPVEYPSLMPNSLAHSNYTLPAETPLALEPS 208
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 233 HSPVRAMP--RGPSALSTSPLLTAPHKLQSGPLVLTEEEKRTLVAEGYPIPT----- 283
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 209 SGVPRAKPTARGEAG-----SRDERRAL-AMKIPPTDKIVNLP 246
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 284 -----KLPLTKSEKALKIRRKKNKISAQESRRKKKKEYMDSLEKKVESCSTENL 334
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 247 "VDDFNELLARYPLTESQALVRDIRRGKGNKVAQAQNCRRKRLTIVQLERELERLTNERE 306
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 335 ELRKKVEVLENTNRTLQLOKL 357
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 307 RLLRARGEADRTLLEVMRQQLTEL 329
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Result No.	Query	Score	Match	Length	DB	ID	Description
1	937.5	35.4	519	4	Q96BA8	Q96ba8 homo sapien	
2	932.5	35.2	519	4	Q96CP0	Q96cp0 homo sapien	
3	928	35.0	507	11	Q92125	Q92125 mus musculus	
4	928	35.0	520	11	Q91W70	Q91w70 mus musculus	
5	436	16.5	516	5	Q24282	Q24282 drosophila	
6	435	16.4	518	5	Q9VU04	Q9vu04 drosophila	
7	394	14.9	395	4	Q8TEV5	Q8tev5 homo sapien	
8	384.5	14.5	461	4	Q96TB9	Q96tb9 homo sapien	
9	378.5	14.3	479	11	Q91XE9	Q91xe9 mus musculus	
10	355	13.4	370	11	Q9D2A5	Q9d2a5 mus musculus	
11	344.5	13.0	315	11	Q9DAE0	Q9dae0 mus musculus	
12	337	12.7	365	6	Q8S019	Q8sg19 bos taurus	
13	330.5	12.5	371	4	Q14671	Q14671 homo sapien	
14	329.5	12.4	371	4	Q9H2W3	Q9h2w3 homo sapien	
15	326.5	12.3	371	4	Q14919	Q14919 homo sapien	
16	326.5	12.3	371	4	Q96GK8	Q96gk8 homo sapien	

```
QY 206 PPSHSDSGSLSPNRLHPFSLQAHSPVPRMPGPSALSTPLTAPHKLQ--SGPL 264
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 208 PSSHSGSDSGSQSPR-SLPP-----SPVPRMARSTAISTPLTAPHKLQTSGLP 260
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 265 VLTEBEKRTLVAEGYPIPTKPLPTKSEKALKKIRKIKNKISQAESRRKKKEYMDSLEK 324
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 261 LUTEBEKRTLIAEGYPIPTKPLPTKAEKALKRVRKIKNKISQAESRRKKKEYVECLEK 320
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 325 KYESCSSTENLELRKKVEVLENTNRTLLQLOKLTLMVGKVSRTCKLAGTGTCTLMVVV 384
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 321 KYVETFTSENELWKVETLENARTLLQLOKLTLMVTKISRPYKMAATQTGTCTLMVAA 380
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 385 LCFVAVFGSFFQGYGYPVPSATKMAIPSOHPLSEP--YTASVVRNLLIYEEHAPLEESS 442
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 381 LCFVLVLGSLVPCLPFSSGQTV--KEDPLAADGVYTASQMPSPRSLLFYDDGA----- 432
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 443 SPASTGELGWDGRGSSLL 460
|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
Db 433 -----GLWEDGRSTL 442
|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
RESULT 2
Q96CPO
ID Q96CPO PRELIMINARY; PRT; 519 AA.
AC Q96CPO;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Similar to old astrocyte specifically induced substance.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC014097; AAH14097.1; -.
DR InterPro; IPR004827; TF_BZIP.
DR Pfam; PF00170; BZIP; 1.
DR PROSITE; PS00036; BZIP_BASIC; UNKNOWN 1.
KW DNA-binding; Nuclear protein.
SQ SEQUENCE 519 AA; 57031 MW; D1313P50FA02A6A8 CRC64;
```

Query Match 35.2%; Score 932.5; DB 4; Length 519;
Best Local Similarity 50.0%; Pred. No. 2.1e-55;
Matches 219; Conservative 52; Mismatches 118; Indels 49; Gaps 14;

```
QY 35 HFSLELDDEFSONVLGQLLSDPPLSEKSESEMEVEPSPPTSAPLIQAHSYSLSEPRQTSP 94
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 42 HFVENMEDFSNDLFSFDDPVLDEKSLDLMELD--SPAPGIAHSYSLSGDSAPQSP 99
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 95 FTHAATSDSFNDEBESEKWLSTPEPSPATIKKEPITEEP--PGLVPSVTLTITAITST- 151
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 100 LVPVKMEDTTQD--AEHGAVALGHLKLSINVKQEQ-SPELPVDPPLAASAAAAAATTT 156
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 152 -----PFKEESPLDMNAGDSSCQTLIPKILPEHEVDQFLNFSKPEASVDQLHLPPT 205
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 157 PLLGLSLPSLR--LPHPQAPGE---MTQLPVKAEPLEVNQFLKVTPE---DLVQMPPT 207
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 206 PPSHSDSGSLSPNRLHPFSLQAHSPVPRMPGPSALSTPLTAPHKLQ--SGPL 264
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 208 PSSHSGSDSGSQSPR-SLPP-----SPVPRMARSTAISTPLTAPHKLQTSGLP 260
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 265 VLTEBEKRTLVAEGYPIPTKPLPTKSEKALKKIRKIKNKISQAESRRKKKEYMDSLEK 324
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 261 LUTEBEKRTLIAEGYPIPTKPLPTKAEKALKRVRKIKNKISQAESRRKKKEYVECLEK 320
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 325 KYESCSSTENLELRKKVEVLENTNRTLLQLOKLTLMVGKVSRTCKLAGTGTCTLMVVV 384
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 321 KYVETFTSENELWKVETLENARTLLQLOKLTLMVTKISRPYKMAATQTGTCTLMVAA 380
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 385 LCPAFAFGSFFQGYGYPVPSATKMAIPSOHPLSEP--YTASVVRNLLIYEEHAPLEESS 442
```

```
Db 381 LCFVLVLGSLVPCLPFSSGQTV--KEDPLAADGVYTASQMPSPRSLLFYDDGA----- 432
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 443 SPASTGELGWDGRGSSLL 460
|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
Db 433 -----GLWEDGRSTL 442
|||:|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:
RESULT 3
Q92125
ID Q92125 PRELIMINARY; PRT; 507 AA.
AC Q92125;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE OASIS protein.
GN OASIS OR OASIS.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=ICR;
RC MEDLINE=99280577; PubMed=10350641;
RA Honma Y., Kanazawa K., Mori T., Tanno Y., Tojo M., Kiyosawa H.,
RA Takeda J., Nikaide T., Tsukamoto T., Yokoya S., Wanaka A.;
RT "Identification of a novel gene, OASIS, which encodes for a putative
RT CREB/AP family transcription factor in the long-term cultured
RT astrocytes and gliotic tissue."
RL Brain Res. Mol. Brain Res. 69:93-103(1999).
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE BZIP FAMILY.
DR EMBL; AB017614; BAA75670.1; -.
DR MGD; MGI:1347062; Oasis.
DR InterPro; IPR001630; Leuzip CREB.
DR InterPro; IPR004827; TF_BZIP.
DR Pfam; PF00170; BZIP; 1.
DR PRINTS; PR00041; LEUZIPPCREB.
DR SMART; SM00338; BRLZ; 1.
DR PROSITE; PS00036; BZIP_BASIC; 1.
KW DNA-binding; Nuclear protein.
SQ SEQUENCE 507 AA; 55428 MW; AAE2F859DAA1D978 CRC64;
```

Query Match 35.0%; Score 928; DB 11; Length 507;
Best Local Similarity 47.9%; Pred. No. 4.1e-55;
Matches 221; Conservative 50; Mismatches 138; Indels 52; Gaps 13;

```
QY 35 HFSLELDDEFSONVLGQLLSDPPLSEKSESEMEVEPSPPTSAPLIQAHSYSLSEPRQTSP 94
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 42 HFVENMEDFSNDLFSFDDPVLDEKSLDLMELD--SPAPGIAHSYSLSGDSAPQSP 99
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 95 FTHAATSDSFNDEBESEKWLSTPEPSPATIKKEPITEEP--PGLVPSVTLTITAITSTP 152
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 100 LVPVKMEDTTQD--VEHGAVALGHLKLSINVKQEQ-SPELPVDPPLAASAAAAAATP 156
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 153 -----FEKEESPLDMNAGDSSCQTLIPKILPEHEVDQFLNFSKPEASVDQLHLPPT 207
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 157 PLLGLSPMLRPLPHQAPGE---MTQLPVKAEPPEMSQFLKVTPE---DLVQMPPT 209
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 208 SSSHSDSGSLSPNRLHPFSLQAHSPVPRMPGPSALSTPLTAPHKLQ--SGPL 266
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 210 SSSHSGSDSGSQSPR-SLPP-----SPVPRMARSTAISTPLTAPHKLQTSGLP 262
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 267 VLTEBEKRTLVAEGYPIPTKPLPTKSEKALKKIRKIKNKISQAESRRKKKEYMDSLEK 326
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 263 LUTEBEKRTLIAEGYPIPTKPLPTKAEKALKRVRKIKNKISQAESRRKKKEYVECLEK 322
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 327 KYESCSSTENLELRKKVEVLENTNRTLLQLOKLTLMVGKVSRTCKLAGTGTCTLMVVV 386
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 323 EYVTSNNELWKVETLENARTLLQLOKLTLMVTKISRPYKMAATQTGTCTLMVAA 382
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 387 FAVAFSGSFFQGYGYPVPSATKMAIPSOHPLSEP--YTASVVRNLLIYEEHAPLEESS 446
```

```
Db 383 FVLVLGSLVPCLPAPFSSGSMVTKEDPIAADSYYAASQMPRSLLFFYDDGA----- 432
Qy 447 TCELGWD--RGSS--IRASGLEALPEVDLPH 476
Db 433 ----GSWEDGRGALLFVEPEGWELKPGGPAQRPDHLRH 469

RESULT 4
Q91W70 PRELIMINARY; PRT; 520 AA.
ID Q91W70
AC Q91W70
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Old ascocyte specifically induced substance.
GN OASIS.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
[1]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary Gland;
RA Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC016447; AAH16447.1; -.
DR MGD; MGI:1347062; Oas1.
DR InterPro; IPR004827; TF_bZIP.
DR Pfam; PF00170; bZIP; 1.
DR PROSITE; PS00036; bZIP_BASIC; UNKNOWN 1.
SQ SEQUENCE 520 AA; 57087 MW; DC2652AA1583A9F2 CRC64;

Query Match 35.0%; Score 928; DB 11; Length 520;
Best Local Similarity 47.9%; Pred. No. 4.3e-55;
Matches 221; Conservative 50; Mismatches 138; Indels 52; Gaps 13;

Qy 35 HFESELLDFSONVLQGLSDPFLSKSMEVEPSPAPLQAEHSYSLSEEPRTQSP 94
Db 42 HFVENMEDFSNDLFSFDDFDVLDKSLDLDWELD--SPAPGQAEHSYSLSGDSAPOS 99
Qy 95 FTHAATSDSFNDEEVESEKWLSTFPFPAATIKKEPITEQP--PGLVPSVTLTITAI 152
Db 100 LVPVKNEDTTQD--VEHGAWALGNKLCSLMWKQEQ-SPELPVDPLAASAMAAAMATP 156
Qy 153 -----FEKEESPLDMNAGDSSCQTLIPKILEPHEVDQFLNFSPEASVDQLHLPPT 207
Db 157 PLGLGLSPNRLPIPHQAPGE---MTQLPVIAEPPEMSQFLKVTPE---DLVQMPPTPP 209
Qy 208 SSHSDSGSLSPNRLPHFSLQASHSPVRAMPGPSALSTSPLLTAPHKLOG-SGPLVL 266
Db 210 SSHGSDSGSQSPR-SLPP-----SSPVRPMARSSTAISTSPLLTAPHKLOGSGPL 262
Qy 267 TEERKRTVAEGYPIPTKLPTKSEKALKKTRKIKKISAQESRRKKKEYMDSLEKKV 326
Db 263 TEERKRTVAEGYPIPTKLPTKABEALKVRKRIKKISAQESRRKKKEYVECLEKKV 322
Qy 327 ESCSTENLELRKRVLENTNRTLQQLQKQLTLMVKVSRRTCKLAGTGTGCLMWWVLC 386
Db 323 EYTTSENELMKVKVETLETANRTLQQLQKQLTMTSKISRYKKAATGTGCLMVAALC 382
Qy 387 FAVAFSGFQGGYPSATKMLPSQHLSEPYTASVVRNLLIYEHEAPLEESSPAS 446
Db 383 FVLVLGSLVPCLPAPFSSGSMVTKEDPIAADSYYAASQMPRSLLFFYDDGA----- 432
Qy 447 TCELGWD--RGSS--IRASGLEALPEVDLPH 476
Db 433 ----GSWEDGRGALLFVEPEGWELKPGGPAQRPDHLRH 469

RESULT 5
Q24282 PRELIMINARY; PRT; 516 AA.
ID Q24282
```

```
AC Q24282;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE CRE-binding transcriptional activator (CREB-A).
GN CREBA OR CREB-A OR CG7450.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92375081; PubMed=1508208;
RA Smolik S.W., Rose T.M., Goodman R.H.;
RT "A cyclic AMP-responsive element-binding transcriptional activator in
RT Drosophila melanogaster, dCREB-A, is a member of the leucine zipper
RT family."
RL Mol. Cell. Biol. 12:4123-4131(1992).
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE bZIP FAMILY.
DR EMBL; M87038; AAA28427.1; -.
DR FlyBase; FBgn0004396; CrebA.
DR InterPro; IPR004827; TF_bZIP.
DR Pfam; PF00170; bZIP; 1.
DR SMART; SM00338; BRLZ; 1.
DR PROSITE; PS00036; bZIP_BASIC; 1.
DR DNA-binding; Nuclear protein.
SQ SEQUENCE 516 AA; 56281 MW; 15BBF2108BE53983 CRC64;

Query Match 16.5%; Score 436; DB 5; Length 516;
Best Local Similarity 33.7%; Pred. No. 1.1e-21;
Matches 142; Conservative 54; Mismatches 117; Indels 108; Gaps 16;

Qy 26 ETEALMYH---THFSELLDFSONVLQGLSDPFLSKSMEVEPSPAPLQAEHS 82
Db 111 DPKCLTFHVPPHATP-----ISRLSSNPALN-----TSVADLTFSSGL 149
Qy 83 YSL-SEEPRTQSPFTHAATSD-----SPNDEEVESEKWL-----STEPFS 122
Db 150 QSLQAHQPHHGSGSHVVVANLEHFQPLHLYDNDCCSSSVSLRDGSMSPDICSDIE 209
Qy 123 ATKKEPITEQPGLVPS-----VTLTITAIPT--FEKEESPLDMNAGDSSCQ 171
Db 210 SAIKDEPMSDSCPASPSTQASSQHQLSLNLAHQSEMLFEPKHCGLLTASSNSN-N 268
Qy 172 TLIPK-----IKLEPHEVDQFLNFSPEASVDQLHLPPTPSS--HSSDSE 215
Db 269 SLIKSQQQQQILQGDNLLMAKWEIKSEKQSTNSSDSKSHAGYGIPLTPSSLSFSD 328
Qy 216 GSLSPNRLHPF-----SLQAH-----SPVR-----AMPGPSALST- 248
Db 329 GNLSEPHLAPLSPNATVSIIVANPAGGESSVRVSRTAASITRSSSGSASGSSSTS 388
Qy 249 -----SPLLTAPHKLOGSGPLVITEERKTLVAEGYPIPTKLPTKSEKALKK 300
Db 389 TTTROPHTPLISSQPK-GSTGTLTTEERKTLVAEGYPIPTKLPTKAEKSLKIRR 447
Qy 301 KIKKISAQESRRKKKEYMDSLEKKVESCSTENLELRKRVLENTNRTLQQLQKQLT 360
Db 448 KIKKISAQESRRKKKEYMDSLEKKVESCSTENLELRKRVLENTNRTLQQLQKQL 507
Qy 361 V 361
Db 508 V 508

RESULT 6
Q9VUQ4 PRELIMINARY; PRT; 518 AA.
ID Q9VUQ4
AC Q9VUQ4;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
```

01-MAR-2002 (TREMELrel. 20, Last annotation update)
DB CRSEA protein.
GN CRSEA OR CG7450.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_taxid=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Calnicker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S.J., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter B.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
RA Ballew R.M., Basu A., Bakendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahle C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,
RA Fostler C., Gabrielian A.C., Garg N.S., Gelbart W.M., Glasser K.,
RA Glöckner A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hootin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalil M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milehina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kianos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RA "The genome sequence of Drosophila melanogaster.";
Science 287:2185-2195(2000).
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE BZIP FAMILY.
DR EMBL; AB003530; AAF49621.1; -.
DR FlyBase; FBgn0004396; Creba.
DR InterPro; IPR001630; Leuzip_CREB.
DR InterPro; IPR004827; TF_BZIP.
DR Pfam; PF00170; bzip; 1.
DR PRINTS; PRO0041; LEUZIPPCCREB.
DR SMART; SM00338; BRLZ; 1.
DR PROSITE; PS00036; BZIP_BASIC; 1.
KW DNA-binding; Nuclear protein.
SQ SEQUENCE 518 AA; 56527 MW; 9461C27F2A87E26 CRC64;

Query Match 16.4%; Score 435; DB 5; Length 518;
Best Local Similarity 34.8%; Pred. No. 1.3e-21;
Matches 138; Conservative 50; Mismatches 112; Indels 96; Gaps 14;

QY 48 LQQLSDPFLSEKSEMEVPSPTSPAPLQAEHSYL-SEEPRTQSPFTHAATSD---- 102
DB 129 ISRLSSNPALN-----TSVADLTRSSGLQLAHQHGGSGSHVVVANLEHF 176
QY 103 ----SFNDEVESEKWL-----STEPSATIKKEPTEEQPPGLVPS----- 141

DB 177 QLPQHLNDNCSSVSLRDGSMSPDICSDIEIDESAIKDBPMSPDSSCPASPTSQASS 236
QY 142 ---VTLTTAISTP--FEKESPLDMNAGDSSCOTLIPK-----IKLEP 181
DB 237 QHQLSLNLAHLQSEMLFPFGCGLLLTASSNSN-NSLTKSQORQOQILGQDNLMAKMEI 295
QY 182 HEVDQFLNFPKPEASVDQLHLPPPTPSS-HSDDSGSLSPNPRLHPF-----SLSQAH- 233
DB 236 KSEKOSTNSSDKSHAGYGLPTPTPSSLPDDSGNLSPEHLFAPLSPNATVSVANP 355
QY 234 -----SPVR-----AMPROFSALST-----SPLLTAPHKLGSGPLV 265
DB 356 ACCESSVRVSRTAASITRSSSGSAGASGSTSTVTTTRQPIHTPLISSQPK-GSTGTLL 414
QY 266 LTEREKRTLVAEGYIPYKPLTKSEKALKIRKKNKISAQESRRKKKYMDSLEKK 325
DB 415 LTEREKRTLVAEGYIPYKPLTKAEKSLKIRKKNKISAQESRRKKKYMDSLEKK 474
QY 326 VESCSTENLELRKKVEVLENTNRTLLQQLQQLV 361
DB 475 VEILTENHDYKKRLEGTNANLLSQLKQALV 510

RESULT 7
Q8TEV5 PRELIMINARY; PRT; 395 AA.
ID Q8TEV5
AC Q8TEV5
DT 01-JUN-2002 (TREMELrel. 21, Created)
DT 01-JUN-2002 (TREMELrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMELrel. 21, Last annotation update)
DE Androgen-induced basic leucine zipper.
GN AIBZIP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21818580; PubMed=11830526;
RA Qi H., Fillion C., Labrie Y., Grenier J., Fournier A., Berger L.,
RA El-Alfy M., Labrie C.;
RT "AIBZIP, a Novel bZIP Gene Located on Chromosome 1q21.3 That Is Highly
RT Expressed in Prostate Tumors and of Which the Expression Is Up-
RT Regulated by Androgens in LNCaP Human Prostate Cancer Cells.";
RL Cancer Res. 62:721-733(2002).
DR EMBL; AF394167; AAL76113.1; -.
SQ SEQUENCE 395 AA; 43432 MW; F3DDF288421AB5F2 CRC64;

Query Match 14.9%; Score 394; DB 4; Length 395;
Best Local Similarity 34.0%; Pred. No. 5.8e-19;
Matches 115; Conservative 49; Mismatches 108; Indels 66; Gaps 12;

QY 135 PGLVPSVTLTITATISTPFKEBESPLDMNAGDSSC--QTLIP-----KIKLEPEV--DQ 186
DB 33 PPPEVPVTRL-----QEQLQGWKSGDGRGCLQSEDEFLKLFIDNEVYVSE 82
QY 187 FLNFPKPEASVDQLHLPPPTPSSHSDS-----EGSLSPNPRLHPFSL 230
DB 83 ASPGSDSGISEDPCF-PDGPAPRATSPMLYEVVYEAALERMOQETGPNVGLISQLD 141
QY 231 QAHSVPVRAMPGRGPSALSTSPLLTAPHKLGSG-----PLVTEBEKRLVA 276
DB 142 Q-WSPAFMVP-DSCWSELFPDHAHLPRAGTVAPVCTTLPCQTLFTDEKRLGQ 199
QY 277 EGYPIPTKPLTKSEKALKIRKKNKISAQESRRKKKYMDSLEKKYVESCSTENLEL 336
DB 200 EGVSLPSHLPLTKAEERVLKVRKKNKQSAQDSRRKKKEYIDGLSESVAAQSAQNEL 259
QY 337 RKKEVLENTNRTLLQQLQQLVLMGVKVBRTCKLAGTGTCTLMVVVLFAVAFGFFQ 396
DB 260 QKKVQLERHNI SLVAQLROLQTL-----IAQTSNKA-AQTSTCVLILLFSLAILPSFS 314
QY 397 GYGYPISATKWLPSQHPLEPVTASVVRNLLIYEE 434

Q91XE9	PRELIMINARY;	PRT;	479 AA.
ID	Q91XE9		
AC	Q91XE9;		
DT	01-DEC-2001 (T:EMBLrel. 19, Created)		
DT	01-DEC-2001 (T:EMBLrel. 19, Last sequence update)		
DT	01-MAR-2002 (T:EMBLrel. 20, Last annotation update)		
DE	Similar to CREB/ATF family transcription factor.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC	Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=LIVER;		
RA	Strausberg R.;		
RL	Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.		
DR	ENBL; BC010786; AAH10786.1; -		
DR	InterPro; IPR004827; TF_BZIP.		
DR	Pfam; PF00170; BZIP; 1.		
DR	PROSITE; PS00036; BZIP BASIC; UNKNOWN 1.		
SQ	SEQUENCE 479 AA; 52145 MW; 948A06FE79DB3779 CRC64;		
Query Match 14.3%; Score 378.5; DB 11; Length 479;			
Best Local Similarity 29.0%; Pred. No. 8.3e-18;			
Matches 157; Conservative 62; Mismatches 153; Indels 169; Gaps			
Qy	1 MEVLSEGEQSVLOWDRKLSLSESPGETEALMYHTHFS--LLDEFQNVGLQLLSPFLS 58		
Db	17 MAPJDSMEVLDLFDRO-----DGLRNVEIAEGWILAREQKVLNDSDEFL- 65		
Qy	59 EKSEMEVEPSPSPAPLQAEHSYLSIEEPRTQSPFTHAATSDSDFNDEVESEKWLST 118		
Db	66 ---NCILGFGSDPS-----SPLWSPADSDSGISEDLPD----- 97		
Qy	119 EFPSATKK---RPI-----TEEPDPLVPSVTITTAISTP-PEKSES-----PLDMN 163		
Db	98 --PQDTPRSGETEPANTVARCHTREQCKGCPSYLPSTPCPEPPRTQVQESSVAIDLDM- 154		
Qy	164 AGDSSCQTLIPKIKLEPHEVDQFLNFSKPEASVDQLHLPPPTPPSHSSDSEGSLSFNP 223		
Db	155 -----WSTDITYPE---EPAGSPSRFNLTVKELL-----SGSGDGL----- 188		
Qy	224 LHPFSLQASHSPVRAMPGRGPSALSTPLITAPHKLQSG---PLVLTEBEKRTLVAEGYP 280		
Db	189 -----QHSLSAASQLGP-----GSGHCQELVITEDEKLLAKEGVT 225		
Qy	281 IPTKLPTKSEKALKIRKIRKIKNKISQAQSRKKKEYMDSLEKKVESCSTENLELRKKV 340		
Db	226 LPTQLPTKYEERVLKIRKIRNKQSAQSRKKKEYIDGLENRMSACTAQHQELQKRV 285		
Qy	341 EVLNTNRTLLQLOKLOTVMGKSVRTCKLAGTGTCTLMVVVLCFAVAFSGFQGYGP 400		
Db	286 LHLEKQNLSLLEQLKHLQALWQSTSKP-----AHAGTCIAVLLLSFAL----- 329		
Qy	401 YPSATKMLPSQHPLE-----PYTASVVRNRLLIIEEHAP-----LSESSPAS 446		
Db	330 -----IILFSIFPNSKNKVDSPGDFVVRVFSRTL---HNHAASRVAPDVTGSEVPGP 380		
Qy	447 TGEIGGWDGSSLLRASSGLEA-----LP-----EVDLPHFLISNET-SLEKSVLL 491		
Db	381 WPDVGTGTHKGPS---SGGLSADWGNFLEIPMLDNLTEELDNSTLVLANSTEDLGRATLL 436		
Qy	492 E 492		
Db	437 D 437		
RESULT 10			
Q9D2A5	PRELIMINARY;	PRT;	370 AA.
ID	Q9D2A5		
AC	Q9D2A5;		
DT	01-JUN-2001 (T:EMBLrel. 17, Created)		
DT	01-JUN-2001 (T:EMBLrel. 17, Last sequence update)		


```
QY 415 LS-----EPYASVVRNLLIYEEHAPLESSSPASTGELGWDGSSLLRASSGLE 467
D 237 FQGOSEARPEYQLHGVISRNILTHEN--VTENLESPV-----LKSLE 278
QY 468 ALPEVDLPFLHISNETSLEKSV 489
D 279 ELPEA--PTTNGSTKHLKMRV 298

RESULT 12
Q8SQ19 PRELIMINARY; PRT; 365 AA.
AC Q8SQ19;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Luman.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN RN SEQUENCE FROM N.A.
RA Budihal P.C., Misra V.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF387035; AAL84006.1; -.
SQ SEQUENCE 365 AA; 40556 MW; 4351DCBA066F9822 CRC64;

Query Match 12.7%; Score 337; DB 6; Length 365;
Best Local Similarity 33.7%; Pred. No. 3.9e-15;
Matches 101; Conservative 44; Mismatches 103; Indels 52; Gaps 9;

QY 198 DQLHLPPTPPS--SHSSDSEGLSPNRLHPFSLQAHSPV-----RAMPRGPSA 245
D 52 DFLSCLPSPPAVLNVFNSDPLVQHD--HYLSQEHVSIDLDSYEKERA----- 102
QY 246 LSTGPLLTAHKLQSGPLVITEEKRITLVAGYPIPTKLPTKSEKALKIRKIRK 305
D 103 -QMTPLRVEEPADQIARLIITEEKRILLEKGLTLPGLPLTKMEEQVLKRVRRKIRK 161
QY 306 ISAOESREKKKEYMDSLKESVSCSTENLELRKKVEVLENTNRTLLQQLKLTVMGKV 365
D 162 KSAOESRRKKKYVGGLESRLKYTAQNLQKLVQLEQNLQSLDQLRLQAMVIQTA 221
QY 366 SRTCKLACTGTCTCLMVVLCFAVAFGFFQGYGYPATKALPSOHPLEPYTASVVR 425
D 222 NK-----ASSSTCVLLVLFSCLLLP-----AMYSDDTRGSLPAEHR-----VL 262
QY 426 SRNLLIYEEHAPLESSSPASTGELGWDGSSLLRASSGL-----EALPEVDLP 475
D 263 SRQLRALPSDEP-PQLPEPALQSEVPKDSLNPQLQASNSCCLFLHLPQAPRAEPLQLP 321

RESULT 13
O14671 PRELIMINARY; PRT; 371 AA.
AC O14671;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Luman.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN RN SEQUENCE FROM N.A.
RA Lu R., Yang P., O'Hare P., Misra V.;
RL MEDLINE=97415590; PubMed=9271389;
RT "Luman, a new member of the CREB/ATF family, binds to herpes simplex
virus VP16-associated host cellular factor.";
```

```
RL Mol. Cell. Biol. 17:5117-5126(1997).
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC EMBL; AF009368; AAB69652.1; -.
DR InterPro; IPR004827; TF_BZIP.
DR Pfam; PF00170; bZIP; 1.
DR SMART; SM00338; BRLZ; 1.
DR PROSITE; PS00036; BZIP_BASIC; 1.
KW DNA-binding; Nuclear protein.
SQ SEQUENCE 371 AA; 41319 MW; 22042E496B92535A CRC64;

Query Match 12.5%; Score 330.5; DB 4; Length 371;
Best Local Similarity 30.7%; Pred. No. 1.1e-14;
Matches 116; Conservative 56; Mismatches 123; Indels 83; Gaps 15;

QY 150 STPEKEESPLDMAGDSSCQTLPKIKLEPHEVDQFLNFSPEKASVDQLHLP-PTPPS 208
D 25 TAPDEAVRAPLDW-----ALP-----LSEVPDWEVDLLCSLLSPPA 62
QY 209 SHSSDSEGLSPNRL--HPFSLQAHSPVAMPRGPSALSTSPLLTAPH-----KLGSG 262
D 63 SLNLS-----SSNCLVHHHTYSLPRETVSMDLESESCRKEGTQMTPOHMEELAEQEA 118
QY 263 PLVLTTEEKRTLVAEGYPIPTKLPTKSEKALKIRKIRKISAOESRRKKKEYMDSL 322
D 119 RLVTDEEKSILLEKGLLPETLPKTEEQILKRVRRKIRKRSQASRRKKKYVGG 178
QY 323 EKKVESCSSTENLELRKKVEVLENTNRTLLQQLKLTVMGKVSRTCKLAGTQTCTCLMV 382
D 179 ESRVLKYTAQNWLQNKVQLLEEQNLSDQLRLQAMVVISNKT-----SSSSTCILV 233
QY 383 VVLCFAVAFGFFQGYGYPATKALPSQH-----PLSEPYTASVVRNLLIY 432
D 234 LLVSFCLLLVP-----AMYSDDTRGSLPAEHLVSRQLRALPSDEPYQLEL----- 279
QY 433 EEHAPLESSSPA-STGELGWDGSS--LLRASSGLEAL-----PEVDLPFLI 479
D 280 ----PALQSEVPKDSLTHQ---WLDGSDCVLQAPNGTSCLLHYMPQAPSAEPPEPDL 332
QY 480 SNETSLEKSVLLELQOHL 497
D 333 SSE-PLCRGPILPQANL 349

RESULT 14
O9H2W3 PRELIMINARY; PRT; 371 AA.
AC O9H2W3;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE CAMP responsive element binding protein 3.
GN CREB3
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN RN SEQUENCE FROM N.A.
RA Ben-Yosef T., Francomano C.A.;
RT "Complete nucleotide sequence and genomic structure of the human CAMP
responsive element binding protein 3 (Luman) gene (CREB3).";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE BZIP FAMILY.
DR EMBL; AF211847; AAG43527.1; -.
DR InterPro; IPR004827; TF_BZIP.
DR Pfam; PF00170; bZIP; 1.
DR SMART; SM00338; BRLZ; 1.
DR PROSITE; PS00036; BZIP_BASIC; 1.
KW DNA-binding; Nuclear protein.
SQ SEQUENCE 371 AA; 41451 MW; 243C4C041B3E1125 CRC64;
```

Query Match 12.4%; Score 329.5; DB 4; Length 371;
Best Local Similarity 29.9%; Pred. No. 1.3e-14;
Matches 117; Conservative 55; Mismatches 122; Indels 97; Gaps 16;

QY 160 LDMNAGGSSCQTLIPKIKLEPHEVDQFLNFPKESVDQLHLP---PTPPSHSSDSE- 215
DB 3 LELDAGDQDLIAFLI-----ESEDLTAPDEA---VRAPLDVALPLSEVPSDWEV 50

QY 216 -----GSLSPNRLHPFSLQA-----HSPVRAMPR-----GPSALSTSLTAPH-- 256
DB 51 DLLCSLLSPASLANLSSNCLVHHDTYSLPRETVSMDLSESCREKGTQMTPOHME 110

QY 257 --KLGSGPLVTEEEKRTLVAEGYPIPTKPLTKSEKALKIRKIKNKISAOESRRKKYMDSL 314
DB 111 ELAEOEIALVLTDSEKSLLEKGLIPETLPTKTEBQILKVRKIRKISAOESRRK 170

QY 315 KKEYMDSLEKVESCTENLELRKVEVLENTNRTLLQQLQKLTLMVGKVSRTCKLAGT 374
DB 171 KKVYVGLSERVLKYTAQNMELQNVLLQQLQKLTLMVGKVSRTCKLAGT 225

QY 375 QTGTCLMVVLCFAVAFSGFFQGYGYPYPSATKMLPSQH-----PLSEPYTASVV 424
DB 226 SSSSTCLVLLVFCLLLP-----AMYSSTRGSLPAEHGVLRSQRLALPSEDYPQLEL- 279

QY 425 RSRNLLIYEHAFLPESSSPA-STGELGCGWDRGSSLLRASSG-----LEALPEVDLPHEL 478
DB 280 -----PALQSEVPKDSHQ---WLDGSDCVLQAGNTSCLLHYMPQAP----- 319

QY 479 ISNETSLE-----KSVLLELQOHL 497
DB 320 -SAEPPLEWFPDLPSEPLCRGPILPQANL 349

RESULT 15
O14919
ID O14919 PRELIMINARY; PRT; 371 AA.

AC O14919;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Basic leucine zipper protein LZIP (CAMP responsive element binding protein 3) (LUMAN).
GN LZIP OR CREB3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID:9606;
RN [1]
SEQUENCE FROM N.A.
Freiman R.N., Herr W.;
Genes Dev. 11:0-0(1997).
RL [2]
RN [2]
RP SEQUENCE FROM N.A.
Ben-Yosef T., Francomano C.A.;
RA "Complete nucleotide sequence and genomic structure of the human CAMP responsive element binding protein 3 (luman) gene (CREB3).";
RT Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=SKIN;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE BZIP FAMILY.
DR EMBL; AF029674; AAB84166.1; -
DR EMBL; AF211848; AAG43528.1; -
DR EMBL; BC010158; AAH10158.1; -
DR InterPro; IPR004827; TF_BZIP.
DR Pfam; PF00170; bZIP; 1.
DR SMART; SM00338; BRLZ; 1.
DR PROSITE; PS00036; BZIP_BASIC; 1.
KW DNA-binding; Nuclear protein.
SQ SEQUENCE 371 AA; 41379 MW; 82152E496B9248EC CRC64;

Query Match 12.3%; Score 326.5; DB 4; Length 371;
Best Local Similarity 30.0%; Pred. No. 2.1e-14;
Matches 115; Conservative 54; Mismatches 121; Indels 93; Gaps 15;

QY 150 STPFKEKESPLDMNAGDSSCQTLIPKIKLEPHEVDQFLNFPKESVDQLHLP-PTPPS 208
DB 25 TAPDEAVRAPLDW-----ALP-----LSEVPSDWEVDLLCSLLSPPA 62

QY 209 SHSSDSSEGLSPNRL--HPFSLQAHSVPVAMPRGPSALSTSLTAPH-----KLGSG 262
DB 63 SINILS-----SSNPLVHHDTYSLPRETVSMDLSESCREKGTQMTPOHMEELAEQIA 118

QY 263 PLVLTTEEEKRTLVAEGYPIPTKPLTKSEKALKIRKIKNKISAOESRRKKYMDSL 322
DB 119 RLVLTDSEKSLLEKGLIPETLPTKTEBQILKVRKIRKISAOESRRKKYVVGGL 178

QY 323 EKKVESCTENLELRKVEVLENTNRTLLQQLQKLTLMVGKVSRTCKLAGTGTCLMV 382
DB 179 ESRVLKYTAQNMELQNVLLQQLQKLTLMVGKVSRTCKLAGTGTCLMV 233

QY 383 VVLCFAVAFSGFFQGYGYPYPSATKMLPSQH-----PLSEPYTASVVRNLLIY 432
DB 234 LLVSPCLLLVP-----AMYSSTRGSLPAEHGVLRSQRLALPSEDYPQLEL----- 279

QY 433 EHAFLPESSSPA-STGELGCGWDRGSSLLRASSG-----LEALPEVDLPHELISNETSLE 486
DB 280 -----PALQSEVPKDSHQ---WLDGSDCVLQAGNTSCLLHYMPQAP-----SAEPPLE 326

QY 487 -----KSVLLELQOHL 497
DB 327 WFPDLPSEPLCRGPILPQANL 349

Search completed: March 5, 2003, 06:23:15
Job time : 98 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 5, 2003, 06:17:28 ; Search time 40 Seconds
(without alignments)
1735.589 Million cell updates/sec

Title: US-09-884-566A-2

Perfect score: 2650

Sequence: 1 MEVLESGQSQVLDKRLSE.....LEGNETLKVELERRVNAVTF 521

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

1 number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```

1: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT.*
23: /SID22/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2650	100.0	521	AAW37934	Mouse CAMP regulat
2	2062	77.8	674	ABG10287	Novel human diagn
3	1518	57.3	993	ABG08669	Novel human diagn
4	937.5	35.4	519	ABG66662	Human fat responsi
5	926	34.9	520	ABG66661	Rat fat responsi
6	887.5	33.5	247	ABG50249	Human transcriptio
7	786	29.7	363	AAAB42970	Human ORFX ORF2734
8	735.5	27.8	479	ABG10527	Novel human diagn
9	520	19.6	144	AAAB41941	Human ORFX ORF1705
10	479	18.1	244	AAAB54258	Human pancreatic c

11	435	16.4	518	22	ABG63292	Drosophila melanog
12	432	16.3	219	22	AAW371382	Human gene 15-enco
13	395	14.9	395	20	AAW39682	Secreted protein 1
14	395	14.9	395	20	AAW01635	Human PS214 derive
15	395	14.9	395	21	AAW00019	Human secreted pro
16	395	14.9	395	21	AAW64650	Human human homolo
17	394	14.9	395	22	AAW93311	Human protein HPI0
18	392	14.8	395	22	AAW67820	Human leucine zipp
19	384.5	14.5	461	23	AAW96664	Human Kail (K1) pr
20	380	14.3	226	23	ABW72303	Rat protein isolat
21	358	13.5	252	22	AAU16574	Human novel secret
22	348	13.1	236	22	AAU16573	Human novel secret
23	348	13.1	239	22	AAU16153	Human novel secret
24	330.5	12.5	371	19	AAW75181	Fragment of human
25	326.5	12.3	371	21	AAW42805	Human ORFX ORF2569
26	326.5	12.3	371	21	AAW73331	HYPM clone 673766
27	273	10.3	123	23	ABP42366	Human ovarian anti
28	249	9.4	299	21	AAW43050	Human ORFX ORF2814
29	249	9.4	299	22	AAW25719	Human protein sequ
30	224	8.5	700	21	AAW83636	Endoplasmic reticu
31	211	8.0	670	21	AAW83635	Endoplasmic reticu
32	181.5	6.8	105	22	AAW91486	Human immune/haema
33	178.5	6.7	450	22	ABG19519	Novel human diagn
34	170.5	6.4	391	22	ABG19691	Novel human diagn
35	166	6.3	547	23	ABP41519	Human ovarian anti
36	162.5	6.1	423	22	ABW92497	Human protein sequ
37	160.5	6.1	172	22	ABW03512	Human musculoskele
38	160	6.0	44	22	AAW64792	Human brain expres
39	160	6.0	44	22	AAW77548	Human bone marrow
40	160	6.0	44	22	AAW37732	Peptide #11769 enc
41	160	6.0	44	23	ABG46575	Human peptide enco
42	158.5	6.0	271	22	AAE00656	Human activating t
43	152	5.7	334	15	AAW53644	c-jun gene gene pr
44	152	5.7	334	23	ABW57355	Mouse ischaemic co
45	150	5.7	589	21	AAW26762	Human Nrf2 protein

ALIGNMENTS

RESULT 1

AAW37934
ID AAW37934 standard; Protein; 521 AA.

XX AC AAW37934;

XX DT 10-AUG-1998 (first entry)

XX DE Mouse CAMP regulatory element binding protein.

XX KW Mouse CAMP regulatory element binding; mCREBa; antibody; inhibition;
mCREBa modulator.

XX OS Mus sp.

XX PN W09812910-A2.

XX PD 02-APR-1998.

XX PF 26-SEP-1997; 97WO-US17288.

XX PR 27-SEP-1996; 96US-0721684.

XX PA (ICOS-) ICOS CORP.

XX PI Keegan KS;

XX DR WPI; 1998-230326/20.

XX N-PSDB; AAV29204.

PT Murine CAMP regulatory element binding protein A, mCREBa - useful
for systematic analysis of structure and function of mCREBa and
identification of those molecules with which it will react

XX Claim 1; Pages 30-31; 41pp; English.

PS This is the amino acid sequence of the mouse cAMP regulatory element

CC binding (mCREBA) protein. The products of the mCREBA such as

CC antibodies can be used for the recombinant production of the protein,

CC to identify novel genes encoding binding partner polypeptides for

CC mCREBA, for the systematic analysis of the structure and function of

CC mCREBA and identification of those molecules with which it will react,

CC to identify inhibitors of mCREBA binding to other natural binding

CC partners and to generate rodents that fail to express a functional

CC mCREBA or express a variant mCREBA, useful as models for studying the

CC activities of mCREBA and mCREBA modulators in vivo.

XX Sequence 521 AA;

SQ Query Match 100.0%; Score 2650; DB 19; Length 521;

Best Local Similarity 100.0%; Pred. No. 9.8e-190;

Matches 521; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MEVLESGEVSQVLMQDKLSLSPGTEALMYHTHSELDDFESQNVLGQLSDPFLSEK 60

1 MEVLESGEVSQVLMQDKLSLSPGTEALMYHTHSELDDFESQNVLGQLSDPFLSEK 60

61 SESMEVEPPTSPAPLQAEHSYSLSEEPRTQSPFTHAATSDSFNDEEVESEKWLSTEF 120

61 SESMEVEPPTSPAPLQAEHSYSLSEEPRTQSPFTHAATSDSFNDEEVESEKWLSTEF 120

121 PSATIKKEPITEPQPPGLVPSVLTITTAISTPPEKESPLDMNAGDSSCQTLIPKIKLE 180

121 PSATIKKEPITEPQPPGLVPSVLTITTAISTPPEKESPLDMNAGDSSCQTLIPKIKLE 180

181 PHEVDQFLNFSPEASVDQLHLPTTPSSHSDSEGLSPNPLHPSLSQAHSVRAMP 240

181 PHEVDQFLNFSPEASVDQLHLPTTPSSHSDSEGLSPNPLHPSLSQAHSVRAMP 240

241 RGPALSTSPLLTAPHKLGSGPLVLTTEBKRTLVAEGYPIPTKLPTKSEKALKKIRR 300

241 RGPALSTSPLLTAPHKLGSGPLVLTTEBKRTLVAEGYPIPTKLPTKSEKALKKIRR 300

301 KIKNKISAQSRKKKYEYMDSEKVEKSCSTENLEIRKKVEVLENTNRTLLQQLQTL 360

301 KIKNKISAQSRKKKYEYMDSEKVEKSCSTENLEIRKKVEVLENTNRTLLQQLQTL 360

361 VMKVRSTCKLAGTGTCTCLMVVLCFAVAFSGFFQGYGYPGATKVALPSOHPLEPYT 420

361 VMKVRSTCKLAGTGTCTCLMVVLCFAVAFSGFFQGYGYPGATKVALPSOHPLEPYT 420

421 ASVVRGNLLIYEHAFLPSESSPASTGELGWDGSSLLRASSGLEALPEVDLPHPFLIS 480

421 ASVVRGNLLIYEHAFLPSESSPASTGELGWDGSSLLRASSGLEALPEVDLPHPFLIS 480

481 NETSLEKSVLLELQOHLVSSKLEGNETLKVVLELERRVNAIF 521

481 NETSLEKSVLLELQOHLVSSKLEGNETLKVVLELERRVNAIF 521

RESULT 2

ABG10287

ID ABG10287 standard; Protein; 674 AA.

XX

AC ABG10287;

XX

DT 13-FEB-2002 (first entry)

XX

DE Novel human diagnostic protein #10278.

XX

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder.

XX

OS Homo sapiens.

XX

PN WO200175047-A2.

XX 11-OCT-2001.

XX

PF 30-MAR-2001; 2001WO-US08631.

XX

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

DR N-PSDB; AAS74474.

XX

PT New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess

PT biodiversity -

XX

Claim 20; SEQ ID No 40646; 103pp; English.

XX

CC The invention relates to isolated polynucleotide (I) and

CC polypeptide (II) sequences. (I) is useful as hybridisation probes,

CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome

CC and gene mapping, and in recombinant production of (II). The

CC polynucleotides are also used in diagnostics as expressed sequence tags

CC for identifying expressed genes. (I) is useful in gene therapy techniques

CC to restore normal activity of (II) or to treat disease states involving

CC (II). (II) is useful for generating antibodies against it, detecting or

CC quantitating a polypeptide in tissue, as molecular weight markers and as

CC a food supplement. (II) and its binding partners are useful in medical

CC imaging of sites expressing (II). (I) and (II) are useful for treating

CC disorders involving aberrant protein expression or biological activity.

CC The polypeptide and polynucleotide sequences have applications in

CC diagnostics, forensics, gene mapping, identification of mutations

CC responsible for genetic disorders or other traits to assess biodiversity

CC and to produce other types of data and products dependent on DNA and

CC amino acid sequences. ABG00010-ABG30377 represent novel human

CC diagnostic amino acid sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

XX

SQ Sequence 674 AA;

Query Match 77.8%; Score 2062; DB 22; Length 674;

Best Local Similarity 63.5%; Pred. No. 1.3e-145;

Matches 428; Conservative 16; Mismatches 34; Indels 196; Gaps 4;

QY 28 EALMYHTHSELDDFESQNVLGQLSDPFLSEKSEMEVEPPTSPAPLQAEHSYSLSE 87

DB 2 EALMYHTHSELDDFESQNVLGQLSDPFLSEKSEMEVEPPTSPAPLQAEHSYSLSE 61

QY 88 EPRTQPPFTHAATSDSFNDEEVESEKWLSTDPPTSIKTEPITDPPGLVPSVLTIT 147

DB 62 EPRAQPPFTHITSDSFNDEEVESEKWLSTDPPTSIKTEPITDPPGLVPSVLTIT 121

QY 148 AISTPPEKESPLDMNAGG-DSSCQTLIPKIKLEPHEVDQFLNFSPEAS----- 196

DB 122 AISTPPEKESPLDMNAGG-DSSCQTLIPKIKLEPHEVDQFLNFSPEAS----- 181

QY 197 ----- 196

DB 182 CTCGSSGNSAIRIILAYSLALRVPTETTKKPRDLFGKTGREKKEET'SREKLYVAERKA 241

QY 197 -VDQLHLPPTPPSSHSDSEGLSPNPLHPSLSQAHSVRAMPGRPSALSTSPILITAP 255

DB 242 PVDHLHLPPTPPSSHSDSEGLSPNPLHPSLSQAHSVRAMPGRPSALSTSPILITAP 301

QY 256 HKLQGGPLVLTTEBKRTLVAEGYPIPTKLPLTKSEKALKKIRRKIKNKISAQSRKK 315

DB 302 HKLQGGPLVLTTEBKRTLVAEGYPIPTKLPLTKSEKALKKIRRKIKNKISAQSRKK 361

Qy 316 KEYMDSLEKVKVESCSTENLELRKKVLENTN----- 347
 Db 362 KEYMDSLEKVKVESCSTENLELRKKVLENTNASHGELVIGVGKRRWDPLLTSPHAG 421
 Qy 348 ----- 347
 Db 422 DQCLGSVOGWAGETPRQLDGSRTTLISEEDPGSPMGGEALGAIVANWGREALHAS 481
 Qy 348 -----RTLQLOKLOKLOTLVMGKVSRTCKLA 372
 Db 482 RVGAVIQAPIAAGSDICLSDIRESATSSQTNISVTRTLTLQLOKLOKLOTLVMGKVSRTCKLA 541
 Qy 373 GTQTCGLMVVVLCAVAFSGFQGYGYPSPATKMLPSQHLSPYATASVRSNLLIY 432
 Db 542 GTQTCGLMVVVLCAVAFSGFQGYGYPSPATKMLPSQHLSPYATASVRSNLLIY 601
 Qy 433 BEHAPLEBSSSPASTGELGMDRGSLRLASSGLEALPEVDLPHELISNETSLEKSVILLE 492
 602 BEHSPPESSSPGAGELGGWDRGSLRLV-SGLSRPDVDLPHELISNETSLEKSVILLE 660
 Qy 493 LQOHLVSKLEGNE 506
 Db 661 LQOHLVSAKLEGNE 674

RESULT 3

ABG08669
 ID ABG08669 standard; Protein; 993 AA.

AC ABG08669;

DT 13-FEB-2002 (first entry)

DE Novel human diagnostic protein #8660.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PF 31-MAR-2000; 2000US-0540217.

PF 23-AUG-2000; 2000US-0649167.

PA (HYSEQ-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

PI WPI, 2001-639362/73.

DR N-PSDB; AAS72856.

PT New isolated polynucleotide and encoded polypeptides, useful in
 diagnostics, forensics, gene mapping, identification of mutations
 responsible for genetic disorders or other traits and to assess
 biodiversity -

PS Claim 20; SEQ ID No 39028; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and
 polypeptide (II) sequences. (I) is useful as hybridisation probes,
 polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 and gene mapping, and in recombinant production of (II). The
 polynucleotides are also used in diagnostics as expressed sequence tags
 for identifying expressed genes. (I) is useful in gene therapy techniques
 to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 quantitating a polypeptide in tissue, as molecular weight markers and as

CC a food supplement. (II) and its binding partners are useful in medical
 imaging of sites expressing (II). (I) and (II) are useful for treating
 disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 diagnostics, forensics, gene mapping, identification of mutations
 responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 993 AA;

Query Match 57.3%; Score 1518; DB 22; Length 993;

Best Local Similarity 71.1%; Pred. No. 1.1e-104;

Matches 317; Conservative 24; Mismatches 55; Indels 50; Gaps 6;

Qy 75 PLI-----QAHSYSLSEPRTQSPFTHAATSDSFENDESEKHWYLTSTEPSATIKKEP 129

Db 8 PIITLSHEVEQAMSLCVHPKT-----IVSQDEVESEKHWYLTSTDFSTSIKTEP 56

Qy 130 ITEQPPGLVPSVTLTITAIPTPEKEESPLDMNAGGSSCOTLPKLEPHEVDQFLN 189

Db 57 ITDEPPGLVPSVTLTITAIPTPEKEEPPLEMTGVDSQCOTIIPKLEPHEVDQFLN 116

Qy 190 FSPKEASVDQLHLPTPPSSHSSDSEGLSPNRLHPFLSQAHSPVRAMPSPSALSTS 249

Db 117 FSPKEAPVDHLHLPTPPSSHSGSDSEGLSPNRLHPFLSQAHSPVRAMPSPSALSSS 176

Qy 250 PLLTAPHKLGSGPLVLTBEEKRTLVAEGYIPYTKLPTKSEKALKKIRKIKKISAQ 309

Db 177 PLLTAPHKLGSGPLVLTBEEKRTLVAEGYIPYTKLPTKSEKALKKIRKIKKISAQ 236

Qy 310 ESRRKKKYMDSLKKEKVCSTENLELRKKEVLENTNRTLLQLOKLOKLOTLVMGKVSRTC 369

Db 237 ESRRKKKYMDSLKKEKVCSTENLELRKKEVLENTNRTLLQLOKLOKLOTLVMGKVSRTC 296

Qy 370 KLAGTQTCGLMVVVLCAVAFSGFQGYGYPSPATKMLPSQHLSPYATASVRSNRL 429

Db 297 KLAGTQTCGLMVVVLCAVAFSGFQGYGYPSPATKMLPSQHLSPYATASVEN----- 352

Qy 430 LIYEHAPLEBSSSPASTGELGGWDRGSLRLASSGLE-----ALP----- 470

Db 353 -LEEMDKFLDKSTLPRLNQE-----EDESLENRQVTGSIEATINSLETKSGSPDGFDTDK 406

Qy 471 -----EVDLPHELISNETSLEKSVLL 491

Db 407 FCQRYKEELVPFLKLFOSTEREGIL 432

RESULT 4

ABG66662

ID ABG66662 standard; Protein; 519 AA.

XX AC ABG66662;

XX AC ABG66662;

DT 29-AUG-2002 (first entry)

XX Human fat responsive transcription factor FTF1.

XX Fatty acid regulated gene; polyunsaturated fatty acid disorder;
 PUPA disorder; eczema; cardiovascular disorder; hypertriglyceridaemia;
 dyslipidaemia; atherosclerosis; coronary artery disease;
 KW cerebrovascular disease; peripheral vascular disease; inflammation;
 KW sinusitis; asthma; pancreatitis; osteoarthritis; rheumatoid arthritis;
 KW acne; body weight disorder; obesity; cachexia; anorexia;
 KW psychiatric disorder; cancer; cystic fibrosis; pre-menstrual syndrome;
 KW diabetes; diabetic complication; genetic polymorphism.

XX Homo sapiens.

XX

Query Match 34.9%; Score 926; DB 23; Length 520;
 Best Local Similarity 47.6%; Pred. No. 8.e-61;
 Matches 219; Conservative 49; Mismatches 142; Indels 50; Gaps 11;

QY 35 HFSELLDFSONVQLQSDPFLSEKSEMEVEPSPPTSPAPLIOAEHSYSLSEEPRTQSP 94
 DB 42 HFVENMEDFNDLFSFDFDVLDEKSPLLDMELD--SPAPGIOAEHSYSLSGDSAPQSP 99
 QY 95 FTHAATSDFNDEEVESEKWKYLSTFPPSNTIKKEPITEOPGLVPSVTLTITASTP-- 152
 DB 100 LVFVVKMEDTTQD--MEHGAWALGNKLCSTMVKQEQSPFLPVDPLAASSAMAATWATPPL 157
 QY 153 ---PEKEESPIDMAGGSSCOTLPIKLEPHEVDQFLNFPKESVNDQLHLPTPTPSS 209
 DB 158 LGLSPISRLPIPHQAPGE---WTQLPVKAEPPEMSQFL----KVTQEDLVQMPPTPSS 210
 QY 210 HSDSEGLSNPRLHPFSLSOAHSPVPRAMPGRGPSALSTSPILTAPHKLQG--SGPLVLTE 268
 DB 211 HGSDDSGSQSPR-SLPP-----SSPVRPMARSSTAISTSPILTAPHKLQGTSGPLLLTE 263
 QY 269 EKRTLVAGGYPIPTKLPITKSEBKALKIRKIKNKISAQSRKKKKEYMDSLEKKVES 328
 DB 264 EKRTLVAGGYPIPTKLPITKAEKALKVRKIRKIKISAQSRKKKKEYVECLEKKVET 323
 QY 329 CSTENLELRKKVELENTNRTLLQOLQKLOTLVMGKVSRCTCKLACTGTCTCLMVVLGFA 388
 DB 324 YTSNENLEWKVETLETANRTLLQOLQKLOTLVTSKISRPYKMAATQTGTCLMVVAALCFV 383
 QY 389 VAFGSFFQGYGYPSPATKMAPLSQHPLSEPYTASVVRGNLLIYEHAFLPESSSPASTG 448
 DB 384 LVGLSLAPCLPAFSGSKTVKEDPVAADSVYAASQMPSSLLFYDDGA-----431
 QY 449 ELGHWG---RSSL-----LRASSGLEALPEVDLPH 476
 DB 432 --GSWEDHGRGALLVEPPEGWELKPGGPAEPRPQDHLRH 469

RESULT 6
 ABB50249
 ID ABB50249 standard; Protein; 247 AA.
 XX ABB50249;
 XX
 DT 05-FEB-2002 (first entry)
 DE
 DE Human transcription factor TRFX-100.
 XX
 XX Human; transcription factor; TRFX; cell proliferative disease;
 XX autoimmune disease; inflammation; neurological disease;
 XX developmental disorder; cancer; AIDS; infection; cytostatic; anti-HIV;
 KW neuroprotective; antiinflammatory; gene therapy.
 KW
 OS Homo sapiens.
 XX
 XX WO200172777-A2.
 XX
 PD 04-OCT-2001.
 XX
 XX 13-MAR-2001; 2001WO-US08117.
 XX
 PR 13-MAR-2000; 2000US-0188986.
 XX
 XX (INCY-) INCYTE GENOMICS INC.
 XX
 XX Hillman JL, Baughn MR, Yue H, Lal P, Lu DAM, Patterson C;
 PI Azimzai Y, Bandman O, Tang YT, Mathur P, Shah P, Au-Young J;
 PI Reddy R;
 XX
 XX WPI; 2001-570896/64.
 DR N-PSDB; ABA83073.
 XX
 XX Novel transcription factor polypeptides, used to treat diseases
 PT associated with altered activity and expression of TRFX, and to screen

PT for agents capable of modulating its activity -
 XX Claim 1; Page 244; 327pp; English.
 XX

The present sequence is the protein sequence for a human transcription factor. The transcription factor and its coding sequence are useful in the diagnosis, treatment and prevention of diseases associated with altered expression of the transcription factor e.g. cell proliferative, autoimmune/inflammatory, neurological and developmental disorders. A number of specific disorders/diseases are given in the specification, including: arteriosclerosis, cirrhosis, hepatitis, cancers, AIDS, allergies, anaemia, asthma, autoimmune thyroiditis, bronchitis, atopic dermatitis, diabetes mellitus, emphysema, osteoarthritis, pancreatitis, Grave's disease, multiple sclerosis, osteoarthritis, psoriasis, rheumatoid arthritis, systemic lupus erythematosus, ulcerative colitis, uveitis, Alzheimer's disease, Huntington's disease, Parkinson's disease, stroke, and viral, bacterial, fungal and protozoal infections.

XX Sequence 247 AA;

Query Match 33.5%; Score 887.5; DB 22; Length 247;
 Best Local Similarity 88.7%; Pred. No. 2.4e-58;
 Matches 172; Conservative 8; Mismatches 13; Indels 1; Gaps 1;

QY 1 MEVLESGEQSVLOWDRKLSLSESGCTEALMYHTHFSELLDFSONVQLQSDPFLSEK 60

DB 1 MEVLESGQGVLOWDRKLSLSESGDGEALMYHTHFSELLDFSONVQLQSDPFLSEK 60

QY 61 SESMEVEPSPSPAPLIOAEHSYSLSEEPRTQSPPTHAATSDSFNDEVESEKWLSTEF 120

DB 61 SVSMEVEPSPSPAPLIOAEHSYSLCEPRAQSPPTH-ITSDSFNDEVESEKWLSTDF 119

QY 121 PSATIKKEPITEOPGLVPSVTLTITAITPPEKEESPLDMAGGSSCOTLPIPKILE 180

DB 120 PSTSIKTEPITDEPPGLVPSVTLTITAITPPEKEEPPLEMTGVDSCCOTIIPKILE 179

QY 181 PHEVDQFLNFPKE 194

DB 180 PHEVDQFLNFPKE 193

RESULT 7

AAAB42970
 ID AAB42970 standard; Protein; 363 AA.
 XX AAB42970;
 XX

DT 08-FEB-2001 (first entry)

DE Human ORFX ORF2734 polypeptide sequence SEQ ID NO:5468.

XX Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
 KW vulnary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiac;
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
 KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
 KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
 KW cardiovascular disease; diabetes mellitus; erythroidism; SCID; AIDS;
 KW cholesterol ester storage; systemic lupus erythematosus; infection;
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
 KW thrombosis; contraceptive.

XX Homo sapiens.

OS WO200058473-A2.

XX

XX 05-OCT-2000.

XX 31-MAR-2000; 2000WO-US08621.

```
XX 31-MAR-1999; 99US-0127607.
PR 02-APR-1999; 99US-0127636.
PR 05-APR-1999; 99US-0127728.
PR 30-MAR-2000; 2000US-0540763.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Shinketsu RA, Leach M;
XX
XX WPI; 2000-602362/57.
XX N-PSDB; AAC77179.
XX
XX Novel nucleic acids and peptides derived from open reading frame X,
XX useful for treating e.g. cancers, proliferative disorders,
XX neurodegenerative disorders and cardiovascular disease -
XX
XX Claim 11; Page 4646-4647; 5507pp; English.
XX
XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
XX which represent the human ORFX open reading frames 1 to 3161. The ORFX
XX sequences have activities such as: cytostatic; hepatotropic; vulnery;
XX antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
XX osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
XX immunostimulant; cardiac; thrombolytic; coagulant; vasotropic;
XX antidiabetic; hypotensive; dermatological; immunosuppressive;
XX antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
XX antithyroid; and antianemic. The sequences can be used for determining
XX the presence of or predisposition to, or preventing or treating
XX pathological conditions associated with an ORFX-associated disorder. The
XX nucleic acids can be used to express ORFX proteins in gene therapy
XX vectors. The proteins and nucleic acids may be used to treat cancers,
XX proliferative disorders, neurodegenerative disorders, osteoarthritis,
XX graft vs host disease, cardiovascular disease, diabetes mellitus,
XX hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
XX erythematosus, severe combined immunodeficiency (SCID), AIDS, viral
XX bacterial or fungal infection, malaria, autoimmune disorders, asthma,
XX allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
XX nocturnal haemoglobinuria, antiinflammatory disease; to enhance
XX coagulation; to inhibit thrombosis; and as a contraceptive.
XX
XX Sequence 363 AA;
XX
XX Query Match 29.7%; Score 786; DB 21; Length 363;
XX Best Local Similarity 52.8%; Pred. No. 1.6e-50;
XX Matches 187; Conservative 34; Mismatches 91; Indels 42; Gaps 13;
XX
XX 35 HFSELDDEFQNVGLQDLPFLSEKSESEVEPSPSPAPLIQAHSYSLSEPRTPSP 94
XX |||:|||||:|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX 42 HFTENMEDFSNLDLFSFDFDVLDEKSPLDLMELD--SPTPGIQAHSYSLSGDSAPQSP 99
XX
XX 95 FTHAATSDSFNDEVESEKWLSTPEPSPATIKKEPITEOP--PGLVPSVTLTITAITST- 151
XX |||:|||||:|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX 100 LVPIKMDTTQD--AEHGAWALGHKLCSINVQEQ-SPELPVDPPLAAPSAMAAAAWATT 156
XX |||:|||||:|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX 152 -----PFEKEEPLDMNAGDSSCQTLPKIKLEPHEVDQFLNFSKEASVDQLHLPPT 205
XX |||:|||||:|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX 157 PLLGLSLPSLR--LPIPHQAPGE---MTQLPVIKAEPLVNVQFLKVTPE---DLVQMPPT 207
XX |||:|||||:|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX 206 PPSHSDSDSGSLSPNRLPFLSPFSLQAHSVPVRAMPGRPSALSTSPLTATPHKLG--SGPL 264
XX |||:|||||:|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX 208 PPSHSDSDSGSOSPR--SLPP-----SSPVRPMARSTAISSPLTATPHKLGSTGSP 260
XX |||:|||||:|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX 265 VLTEEEKRTLVAGGYPTKPLTKSEKALKKIRRIKKNKISQAESRRKKKKEYMDSLEK 324
XX |||:|||||:|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX 261 VLTEEEKRTLVAGGYPTKPLTKSEKALKKIRRIKKNKISQAESRRKKKKEYMDSLEK 320
XX |||:|||||:|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX 325 KVESCSTENLELRKKVEVLNNTNLTLLQQLQKQLVMGVKVRCKLAGTGTGT 378
XX *|||||:|||||:|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX 321 KVESCSTENLELRKKVETLENAN-SFSSGIQPL-----LCSLTIGLENPT 363
XX |||:|||||:|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX
XX RESULT 8
```

```
ABG10527
ID ABG10527 standard; Protein; 479 AA.
XX
XX AC ABG10527;
XX
XX DT 13-FEB-2002 (first entry)
XX
XX DE Novel human diagnostic protein #10518.
XX
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder.
XX
XX OS Homo sapiens.
XX
XX PN WO200175067-A2.
XX
XX PD 11-OCT-2001.
XX
XX PF 30-MAR-2001; 2001WO-US08631.
XX
XX PR 31-MAR-2000; 2000US-0540217.
XX
XX PR 23-AUG-2000; 2000US-0649167.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
XX N-PSDB; AAS74714.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity -
XX
XX Claim 20; SEQ ID No 40886; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX and gene mapping, and in recombinant production of (II). The
XX polynucleotides are also used in diagnostics as expressed sequence tags
XX for identifying expressed genes. (I) is useful in gene therapy techniques
XX to restore normal activity of (II) or to treat disease states involving
XX (III). (II) is useful for generating antibodies against it, detecting or
XX quantitating a polypeptide in tissue, as molecular weight markers and as
XX a food supplement. (II) and its binding partners are useful in medical
XX imaging of sites expressing (II). (I) and (II) are useful for treating
XX disorders involving aberrant protein expression or biological activity.
XX The polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. ABG00010-ABG30377 represent novel human
XX diagnostic amino acid sequences of the invention.
XX Note: The sequence data for this patent did not appear in the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 479 AA;
XX
XX Query Match 27.8%; Score 735.5; DB 22; Length 479;
XX Best Local Similarity 43.7%; Pred. No. 1.4e-46;
XX Matches 194; Conservative 48; Mismatches 95; Indels 107; Gaps 18;
XX
XX 63 SMEVEPSPSPAPLIQAHSYSLSEPRTPSPPTTHAATSD-SFND----- 106
XX |||:|||||:|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX 6 AMATSPA-AGPQLVQAAH-----PPDSTSIYTPGLISDFDPSGILGAPSLMPPSVL 59
XX |||:|||||:|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX 107 -----EEVESEKWLSTPEPSPATIKKEPITE-----EQPPGL--VPSVTLTITAI 150
XX |||:|||||:|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX 60 LLPAIEQLHCSRHYLL--FYS---HGNPHTKACVAFVGVGGPFLGSPSLSCAQN 114
XX |||:|||||:|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

QY 151 TPFE-----KEESPLDNAGDSSCOTLIPKIKLEPHEVDOLFSPKEASVDQ 199
 Db 115 KSHFPNCAGQATSPREEAPGEM-----TQLPVIKAEPLVNOFLKVTPE-----DL 161
 QY 200 LHLPTTPSSSHSDSGSLSPNRLHPFSLSOAHSPVRAMPGRGPSALSTSPLLTAPHKIQ 259
 Db 162 VQMPTTPSSSHSGSDSGSPR-SLPP-----SSPVRPMARSSTAISTSPLLTAPHKIQ 214
 QY 260 G-SGLVLTEEEKRTLVAEGYDIPKPLPTKSEBKALKIRKIKNKISQAQSRKKKEY 318
 Db 215 GTSGLPTEEEKRTLVAEGYDIPKPLPTKAEKALKVRRKIKNKISQAQSRKKKEY 274
 QY 319 MDSLEKVESCTENLELRKKVLENTNRTLLQLOQLQTLVGMKVSRCKLAGTQTGT 378
 Db 275 VECLKKVETFSNNELWKK-----LQKLQTLVTKNISRPYKMAATQTGT 320
 QY 379 CLMVVLCFAVAFGFFQGYGYPYKATKMAQPSHPLSEP--YTASVVRNRLLIYEHA 436
 Db 321 CLMVAALCFVLVGLSLVPLCPFFSGQTV--KEDPLAADGVVTASQMPFSLFFYDDGA 378
 QY 437 PLEESSSPASTGELGWDGSGSL 460
 Db 379 -----GLWEDGRSTL 388
 RESULT 9
 AAB41941
 ID AAB41941 standard; Protein; 144 AA.
 XX
 AC AAB41941;
 XX
 DT 08-FEB-2001 (first entry)
 XX
 DE Human ORFX ORF1705 polypeptide sequence SEQ ID NO:3410.
 XX
 KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
 KW vulnary; antipariatic; antiparkinsonian; nootropic; neuroprotective;
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiac;
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 KW hypotensive; dermatological; immunosuppressive; antihypertensive;
 KW antiviral; antibacterial; antifungal; antitumoral; antithyroid;
 KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KW cholesterol ester storage; systemic lupus erythematosus; infection;
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
 KW thrombosis; contraceptive.
 XX
 OS Homo sapiens.
 XX
 PN WO200058473-A2.
 XX
 PD 05-OCT-2000.
 XX
 PF 31-MAR-2000; 2000WO-US08621.
 XX
 PR 31-MAR-1999; 99US-0127607.
 PR 02-APR-1999; 99US-0127636.
 PR 05-APR-1999; 99US-0127728.
 PR 30-MAR-2000; 2000US-0540763.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Shimkets RA, Leach M;
 XX
 DR WPI; 2000-602362/57.
 DR N-ESDB; AAC76150.
 XX
 PT Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease -

XX Claim 11; Page 2587; 5507pp; English.
 PS AACT7446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
 XX which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytostatic; hepatotropic; vulnary;
 CC antipariatic; antiparkinsonian; nootropic; neuroprotective;
 CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
 CC immunostimulant; cardiac; thrombolytic; coagulant; vasotropic;
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;
 CC antiinflammatory; antibacterial; antiviral; antifungal; antitumoral;
 CC antithyroid; and antianemic. The sequences can be used for determining
 CC the presence of or predisposition to, or preventing or treating
 CC pathological conditions associated with an ORFX-associated disorder. The
 CC nucleic acids can be used to express ORFX proteins in gene therapy
 CC vectors. The proteins and nucleic acids may be used to treat cancers,
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
 CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
 CC coagulation; to inhibit thrombosis; and as a contraceptive.
 XX
 SQ Sequence 144 AA;
 Query Match 19.6%; Score 520; DB 21; Length 144;
 Best Local Similarity 89.4%; Pred. No. 3.4e-31;
 Matches 101; Conservative 4; Mismatches 8; Indels 0; Gaps 0;
 QY 1 MEVLESGQSVLQWDRKLSLSEPGETEALMYHFSSELLDEFPSONVLQQLSDPFLSEK 60
 Db 1 MEVLESGQSVLQWDRKLSLSEPGDGEALMYHFSSELLDEFPSONVLQQLSDPFLSEK 60
 QY 61 SESMEVEPSPTSPAPLQAEHSYLSLSEPPRTQSPETHAATSDSFNDEVESEK 113
 Db 61 SVSMEVEPSPTSPAPLQAEHSYLSLSEPPRTQSPETHAATSDSFNDEVESEK 113
 RESULT 10
 AAB54258
 ID AAB54258 standard; Protein; 244 AA.
 XX
 AC AAB54258;
 XX
 DT 09-MAR-2001 (first entry)
 XX
 DE Human pancreatic cancer antigen protein sequence SEQ ID NO:710.
 XX
 KW Human; pancreas; pancreatic cancer; pancreatic cancer antigen;
 KW detection; diagnosis; identification; cytostatic; neuroprotective;
 KW nootropic; immunomodulatory; relaxant; contraceptive; gynaecological;
 KW antiinflammatory; cardiac; gene therapy; chromosome mapping;
 KW linkage analysis; tissue identification; tissue typing; forensic;
 KW neural; immune system; muscular; reproductive; gastrointestinal;
 KW pulmonary; cardiovascular; renal; proliferative.
 XX
 OS Homo sapiens.
 XX
 PN WO200055320-A1.
 XX
 PD 21-SEP-2000.
 XX
 PF 08-MAR-2000; 2000WO-US05989.
 XX
 PR 12-MAR-1999; 99US-0124270.
 PR (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM;
 XX
 DR WPI; 2000-579444/54.

DR N-PSDB; AAC99023.
XX New nucleic acid that is a pancreatic cancer antigen for preventing,
PT treating, or ameliorating a medical condition, particular pancreatic
PT cancer, or for use in assays for diagnosing a pathological condition -
XX
XX Claim 11; Page 1150-1151; 1379pp; English.
XX
XX AAC98773 to AAC99231 encode the human pancreatic cancer associated
CC proteins, called pancreatic cancer antigens, given in AAB54008 to
CC AAB54466. The human pancreatic cancer antigens have cytostatic,
CC neuroprotective, nootropic, immunomodulatory, relaxant, contraceptive,
CC gynaecological, cardiant and antiinflammatory activities, and can be used
CC in gene therapy. The polynucleotide and proteins can be used for
CC preventing, treating, or ameliorating a medical condition or in assays
CC for diagnosing a pathological condition or a susceptibility to one in a
CC subject. Binding partners to the proteins and the activity of the
CC proteins can be identified. The pancreatic cancer antigens can be used to
CC detect, treat or prevent pancreatic disorders, especially cancer.
CC Agonists and antagonists to the antigens can be screened for. The
CC pancreatic cancer antigen polynucleotides can be used to design nucleic
CC acid hybridisation probes that can be used in chromosome mapping, linkage
CC analysis, tissue identification and/or typing and a variety of forensic
CC and diagnostic methods. The proteins can be used to generate antibodies
CC which are used to purify, detect and target the polypeptides, including
CC both in vivo and in vitro diagnostic and therapeutic methods. The
CC proteins can be used to treat or prevent neural, immune system, muscular,
CC reproductive, gastrointestinal, pulmonary, cardiovascular, renal or
CC proliferative disorders. AAC99232 to AAC99240 and AAB54467 represent
CC sequences used in the exemplification of the present invention.
XX
XX Sequence 244 AA;

Query Match 18.1%; Score 479; DB 21; Length 244;
Best Local Similarity 57.4%; Pred. No. 8.3e-28;
Matches 105; Conservative 21; Mismatches 39; Indels 18; Gaps 3;
QY 280 PIPKPLTKSEKALKKTRKTKNISQESRRKKKEYMDSLEKKVESCSTENLELRKK 339
DB 1 PIPKPLTKAEKALKVRNRKKNKISQESRRKKKEYVECLEKKKVVETPTSENELWKK 60
QY 340 VELENTNRTLLQQLKQLTVLMGVKVSRTCKLAGTGTGTCMLVVLVCFVAFGSPFGYG 399
DB 61 VETLENANRTLLQQLKQLTVLTKNSRKYKMAATGTGTCMLVAALCFVLVLSLVPCLP 120
QY 400 PYPATNMALPSQHPLEP--YTASVVRSNLLIYEHAPLESSESSPASTGELGGWDRGS 457
DB 121 EFSGSGQTV--KEDPLAAGVYTASQMPRSRLFLFYDDGA-----GLWEDGR 164
458 SLL 460
DB 165 STL 167

RESULT 11
AAB63292
ID AAB63292 standard; Protein; 518 AA.
XX AAB63292;
XX
XX 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster polypeptide SEQ ID NO 16668.
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX Drosophila melanogaster.
XX
XX WO200171042-A2.
XX
XX 27-SEP-2001.
XX

PF 23-MAR-2001; 2001WO-US09231.
XX
XX 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
XX (PEKE) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI; 2001-655860/75.
DR N-PSDB; ABL07395.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
XX Disclosure; SEQ ID NO 16668; 2lpp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABLI16176-ABLI30511), expressed DNA
CC sequences (ABLI01840-ABLI16175) and the encoded proteins
CC (AAB57737-ABBI72072).
XX
XX The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 518 AA;

Query Match 16.4%; Score 435; DB 22; Length 518;
Best Local Similarity 34.8%; Pred. No. 4.7e-24;
Matches 138; Conservative 50; Mismatches 112; Indels 96; Gaps 14;
QY 48 LQQLSDPFLSEKSEMEVPSPTGPAPLIQAEHSYSL-SSEPRTOQSPFTHAATSD---- 102
DB 129 ISRLSNPALN-----TSVADLTRSGSLQSLQAHQPHHGGSGSHVVANLEHF 176
QY 103 -----SFNDEVESEKWL-----STFPSTATIKKEPITEEPQPLVPS----- 141
DB 177 QLPQHLVNDNCSSVSLRDGSMSPDICSDIETDESARKDEPMSPDSCPASPTQASS 236
QY 142 ---VTITITAITSTP--FEKEESPLDMNAGDSSCQTLIPK-----TKLEP 181
DB 237 QHQLSLNLAHQSEMULFEPKHCGCLLLTASNEN-NSLKSQORQOILQODNLLMAKMEI 295
QY 182 HEVDQFLNFPKESAVDQLHLPTPPSS-HSSDSEGLSPNPLRHPF-----SLSQAH- 233
DB 296 KSEKQSTSNSSDKSHAHGYGIPLTFFSSILPDDSEGNLSPEHLFAPLSNATVISVANP 355
QY 234 ----SPVR-----AMPRGPSALST-----SPLLTAPHKLGSGPLV 265
DB 356 AGGESSVRVSRRTAASITRSSSGSASASGSGSTSTVTTTRQPIHTPLISSQPK-GSTGLL 414
QY 266 LTBEEKRTLVAEGYPIPTKPLTKSEKALKKIRKKNKISQESRRKKKEYMDSLEKK 325
DB 415 LTBEEKRTLLAEGYIPQKPLPTKAEKSLKIRKKNKISQESRRKKKEYMDQLERR 474
QY 326 VESCSTENLELRKKVLENTNRTLLQQLKQLTV 361
DB 475 VEILVTENHDYKRLGLEETNANLLSOLHKLQALV 510
RESULT 12
AAG71382
ID AAG71382 standard; Protein; 219 AA.
XX AAG71382;
XX
XX 30-JUL-2001 (first entry)
XX

Human gene 15-encoded secreted protein fragment, SEQ ID NO:234.

Human; secreted protein; proliferative disorder; cancer; foetal abnormality; developmental abnormality; haematopoietic disorder; immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis; inflammation; allergy; neurological disorder; Alzheimer's disease; Parkinson's disease; cognitive disorder; schizophrenia; asthma; skin disorder; psoriasis; sepsis; diabetes; atherosclerosis; cardiovascular disorder; angiogenic disorder; kidney disorder; gastrointestinal disorder; pregnancy-related disorder; tumour; endocrine disorder; infection; wound healing; vulnery; cell culture; chemotaxis; food additive; binding partner identification.

Homo sapiens.

WO200132674-A1.

10-MAY-2001.

25-OCT-2000; 2000WO-US293360.

29-OCT-1999; 99US-0162211.

30-JUN-2000; 2000US-0215138.

(HUMA-) HUMAN GENOME SCI INC.

Ruben SM, Komatsoulis GA, Young PE, Moore PA; WPI; 2001-291051/30.

New nucleic acid molecule encoding a human secreted protein, useful for preventing, treating or ameliorating medical conditions such as rheumatoid arthritis, Alzheimer's disease and microbial infections - Disclosure; Page 39; 581pp; English.

AAH31349-AAH31428 represent cDNAs corresponding to 26 human secreted protein genes, and AAG71243-AAG71319 represent the proteins they encode. AAG71320-AAG71403 represent human secreted protein fragments. The genes and their corresponding secreted proteins are useful for preventing, treating or ameliorating medical conditions, e.g., by protein or gene therapy. Pathological conditions can be diagnosed by determining the amount of the new protein in a sample or by determining the presence of mutations in the new genes. Specific uses are described for each of the 52 genes, based on the tissues in which they are most highly expressed, and include developing products for the diagnosis or treatment of proliferative disorders, cancer, tumours, foetal and developmental abnormalities, haematopoietic disorders, diseases of the immune system, AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation, allergies, neurological disorders (e.g., Alzheimer's disease, Parkinson's disease), cognitive disorders, schizophrenia, asthma, skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders, angiogenic disorders, kidney disorders, gastrointestinal disorders, pregnancy-related disorders, endocrine disorders, and infections. The proteins can also be used to aid wound healing and epithelial cell proliferation, to prevent skin aging due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to regenerate tissues, to identify their cognate ligands or binding partners, and in chemotaxis, and can be used as a food additive or preservative to modify storage properties. Antibodies specific for a protein of the invention can be used in alleviating symptoms associated with the disorders mentioned above, and in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked immunosorbent assay (ELISA). The present sequence represents a human secreted protein fragment referred to in the disclosure of the invention.

Sequence 219 AA;

Query Match 16.3%; Score 432; DB 22; Length 219;
Best Local Similarity 54.9%; Pred. No. 2.3e-24;
Matches 96; Conservative 22; Mismatches 39; Indels 18; Gaps 3;

QY 288 TKSEKALKIRIKKIKNKISAESEKRRKKYMDSLKKVESCSTENLELRKKVLENTIN 347
Db 1 SKAEKALKURVKRIRKIKNKISAESEKRRKKYVECLEKRVETPTSENNELEKVKVLENTAN 60
QY 348 RTLLQOLQKLTLMVNGKVRCTCKLAGTGTCTLMVVLVLCFAVAFSGFFQGYGYPSPATKM 407
Db 61 RTLLQOLQKLTLMVNGKVRCTCKLAGTGTCTLMVVLVLCFAVAFSGFFQGYGYPSPATKM 120
QY 408 ALPSOHPLEP--YTASVVRNRNLLIYEHPLEBSSSPASTGELGGWDRGSSLL 460
Db 121 V--KEDPLADGVYATASQMPRSRLFYDDGA-----GLWEDGRSTL 159
RESULT 13
AAI59682
ID AAY59682 standard; Protein; 395 AA.
XX
AC AAY59682;
XX
DT 18-JAN-2000 (first entry)
XX
DE Secreted protein 108-009-5-0-A2-FL.
XX
KW Secreted protein; fingerprint identification technique;
KW chromosome mapping; human; hereditary disease; diagnosis; cancer;
KW hyperlipidaemia; cardiovascular; neurodegenerative disorder; therapy;
KW autoimmune disease; rheumatic disease; embryonic disorder; myopathy;
KW renal injury; amino aciduria; hypoglycaemia; male rat infertility;
KW hypertension.
XX
OS Homo sapiens.
XX
PN WO9940189-A2.
XX
PD 12-AUG-1999.
XX
PT 09-FEB-1999; 99WO-IB00282.
XX
PR 09-FEB-1998; 98US-0074121.
PR 13-APR-1998; 98US-0081563.
PR 10-AUG-1998; 98US-0096116.
PR 04-SEP-1998; 98US-0099273.
XX
PA (GEST) GENSET.
XX
PI Bougueleret L, Duclert A, Dumas Milne Edwards J;
XX
DR WPI; 1999-600966/51.
XX
N-PSDB; AAZ40810.
XX
Extended cDNAs useful for expressing secreted proteins and to obtain specific antibodies -
XX
Claim 10; Page 202-203; 244pp; English.
XX
This sequence represents a human secreted protein of the invention. The extended cDNAs (or genomic DNAs obtainable from them) may be used to prepare PCR primers and probes. These are useful for forensic matching or positive identification by DNA sequencing. They may also be used in alternative fingerprint identification techniques. Antibodies against the proteins encoded by the extended cDNAs are useful in identification of tissue types or cell species, as well as identifying tissue specific soluble proteins. The sequences can be used for chromosome mapping and identification of genes associated with hereditary diseases or drug response. Signal sequences from the cDNAs can be used in construction of secretion vectors. Other sequences derived from the extended cDNAs can be used to clone upstream genomic DNA sequences including promoters. This is in turn useful for identifying proteins that interact with promoter sequences. Some of the proteins may be useful in diagnosing and treating several disorders including, but not limited to: cancer, hyperlipidaemia, cardiovascular and neurodegenerative disorders, autoimmune diseases, and rheumatic diseases, embryonic disorders, hypertension and myopathies, amino acidurias, hypoglycaemia, male rat infertility and myopathies.

XX SQ Sequence 395 AA;

Query Match 14.9%; Score 395; DB 20; Length 395;
 Best Local Similarity 34.0%; Pred. No. 3.1e-21;
 Matches 115; Conservative 49; Mismatches 108; Indels 66; Gaps 12;

Qy 135 PPGLVPSVTLTITATSTPFKEESPLDNAGDSSC--QTLP-----KIKLEPHEV--DQ 186
 Db 33 PPPEVPTL-----QEQGLQGMKSGDRCGLQESPEDFLKLFDIDPNEVYCS 82

Qy 187 FLNFPKESVDQLHLPPTPPSSHSDS-----EGSLSPNRLHPFSL 230
 Db 83 ASPGSDGISDSC--PSPAPRATSSPMLYEVVEAGALERMOETGPNVGLISIQLD 141

Qy 231 QAHSVVRAMPGPSALSTPLTAPHKLGSG-----PLVLTBEERKTLVA 276
 Db 142 Q--WSPAFWVP--DSCMVSELPDPAHAHILPRAGTVAPVCTTLPCQTLFLTDEKRLLGQ 199

Qy 277 EGYPIPTKLPLTKSEKALKIRKIKNKSIAQESRRKKKYMDSLEKKVESCSTENLEL 336
 Db 200 EGVSLPSHLPLTKAEERVLKVRKIRNKQSAQDSRRKKKEYIDGLSRVAACSAQNQEL 259

Qy 337 RKKVEVLENTNLTLLQOLQKLTVMGKVSRTCKLAGTGTGTCMLVWLCFVAVAFSGFFQ 396
 Db 260 QKKVQELERHNLISLVAQLRQLTL-----IAQTSNKA-AQTSTCVLILLPSLAILILPSFS 314

Qy 397 GYGPVPSATKMALPSQHPLSEPTYASVVRNLLIYEE 434
 Db 315 PQSRPEAG-----SEDYQPHGVTSRNILTHKD 342

RESULT 14
 AAY01635
 ID AAY01635 standard; Protein; 395 AA.
 XX AC AAY01635;
 XX DT 22-JUN-1999 (first entry)
 XX DE Human PS214 derived polypeptide.
 XX KW PS214; prostate cancer; breast cancer; ovarian cancer;
 KW benign prostatic hyperplasia; BPH; prostatitis; ovarian cyst;
 KW prostatic intraepithelial neoplasia; PIN; atypical hyperplasia;
 KW fibroadenoma; cystic breast disease; ovarian cyst adenoma;
 KW ovarian endometrioma.
 XX PA (ABBO) ABBOTT LAB.
 XX WO9914357-A2.
 XX PD 25-MAR-1999.
 XX PF 18-SEP-1998; 98WO-US19496.
 XX PR 23-DEC-1997; 97US-0997074.
 XX PR 19-SEP-1997; 97US-0938383.
 XX PA (ABBO) ABBOTT LAB.
 XX PI Billing-Medel PA, Cohen M, Colpitts TL, Friedman PN;
 PI Gordon J, Granados EN, Hodges SC, Klass MR, Kratochvil JD;
 PI Merchant BT, Roberts-Rapp L, Russell JC, Stroupe SD;
 XX PR WPI; 1999-229547/19.
 XX PT New isolated PS214 polynucleotides and polypeptides used for, e.g.
 PT preventing or treating breast cancer
 XX PS Claim 23; Page 113-114; 116pp; English.
 XX PT Amino acid sequences AAY01635-38 represent PS214 derived polypeptides.

CC They are used for detecting, staging, preventing or treating diseases
 CC or conditions of the prostate, breast and ovary, particularly cancer.
 CC The polynucleotides are useful for detecting, diagnosing, staging,
 CC monitoring, prognosticating, in vivo imaging, preventing or treating,
 CC or determining predisposition to diseases or conditions of the prostate,
 CC breast and ovary such as prostate cancer, breast cancer and ovarian
 CC cancer. They can be used for conditions such as e.g. benign prostatic
 CC hyperplasia (BPH), prostatitis, prostatic intraepithelial neoplasia
 CC (PIN), atypical hyperplasia, fibroadenoma, cystic breast disease,
 CC ovarian cysts, ovarian cyst adenoma, or ovarian endometriomas.
 XX SQ Sequence 395 AA;

Query Match 14.9%; Score 395; DB 20; Length 395;
 Best Local Similarity 34.0%; Pred. No. 3.1e-21;
 Matches 115; Conservative 49; Mismatches 108; Indels 66; Gaps 12;

Qy 135 PPGLVPSVTLTITATSTPFKEESPLDNAGDSSC--QTLP-----KIKLEPHEV--DQ 186
 Db 33 PPPEVPTL-----QEQGLQGMKSGDRCGLQESPEDFLKLFDIDPNEVYCS 82

Qy 187 FLNFPKESVDQLHLPPTPPSSHSDS-----EGSLSPNRLHPFSL 230
 Db 83 ASPGSDGISDSC--PSPAPRATSSPMLYEVVEAGALERMOETGPNVGLISIQLD 141

Qy 231 QAHSVVRAMPGPSALSTPLTAPHKLGSG-----PLVLTBEERKTLVA 276
 Db 142 Q--WSPAFWVP--DSCMVSELPDPAHAHILPRAGTVAPVCTTLPCQTLFLTDEKRLLGQ 199

Qy 277 EGYPIPTKLPLTKSEKALKIRKIKNKSIAQESRRKKKYMDSLEKKVESCSTENLEL 336
 Db 200 EGVSLPSHLPLTKAEERVLKVRKIRNKQSAQDSRRKKKEYIDGLSRVAACSAQNQEL 259

Qy 337 RKKVEVLENTNLTLLQOLQKLTVMGKVSRTCKLAGTGTGTCMLVWLCFVAVAFSGFFQ 396
 Db 260 QKKVQELERHNLISLVAQLRQLTL-----IAQTSNKA-AQTSTCVLILLPSLAILILPSFS 314

Qy 397 GYGPVPSATKMALPSQHPLSEPTYASVVRNLLIYEE 434
 Db 315 PQSRPEAG-----SEDYQPHGVTSRNILTHKD 342

RESULT 15
 AAG00019
 ID AAG00019 standard; Protein; 395 AA.
 XX AC AAG00019;
 XX DT 06-OCT-2000 (first entry)
 XX DE Human secreted protein #7.
 XX KW Human; secreted protein; 5' EST; expressed sequence tag; cDNA isolation;
 KW gene therapy; chromosome mapping.
 XX OS Homo sapiens.
 XX PN EP1033401-A2.
 XX PD 06-SEP-2000.
 XX PF 21-FEB-2000; 2000EP-0200610.
 XX PR 26-FEB-1999; 99US-0122487.
 XX PA (GEST) GENSET.
 XX PI Dumas Milne Edwards J, Duclert A, Giordano J;
 XX WPI; 2000-500381/45.
 XX DR N-PSDB; AAC00016.
 XX PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for

PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
PS Example 19; SEQ ID 14; 71pp + CD-ROM; English.
XX
CC The present sequence is a human secreted protein which shows extensive
CC homology to the bZIP family of transcription factors, and especially to
CC the human human protein. The full length cDNA encoding this protein was
CC obtained from a 5' EST using first and second strand synthesis
CC procedures. 5' ESTs were prepared from total human RNAs or polyA+ RNAs
CC derived from 30 different tissues. EST sequences usually correspond
CC mainly to the 3' untranslated region (UTR) of the mRNA because they are
CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
CC well suited for isolating cDNA sequences derived from the 5' ends of
CC mRNAs and even in those cases where longer cDNA sequences have been
CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
CC mRNAs with intact 5' ends and can therefore be used to obtain full length
CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
CC gene therapy and chromosome mapping procedures. They are used to obtain
CC upstream regulatory sequences and to design expression and secretion
CC vectors.
XX
SQ Sequence 395 AA;
Query Match 14.9%; Score 395; DB 21; Length 395;
Best Local Similarity 34.0%; Pred. No. 3.1e-21;
Matches 115; Conservative 49; Mismatches 108; Indels 66; Gaps 12;
QY 135 PPGILVPSVTLTITAITPFKEEESPLDMNAGGDSGC--QTLIP----KIKLEPHEV--DQ 186
Db 33 PPPEVPVTRL-----QEQLGQWKSGGDRCCGLQESEPEDFLKFIDPNEVCSE 82
QY 187 FLNFSPKASVDQLHLPTPPSSHSDS-----EGSLSPNPRHLHPFSL 230
Db 83 ASPGSDSGISEDSCH-PDSPAPRATSPMLYEVVYEAALERMQGETGPNVGLISIQLD 141
QY 231 QAHSVPVRMPRGPSALSTSLTAPHKLGSG-----PLVITEEKRTIVA 276
Db 142 Q-WSPAFWVP--DSCWVSELPPDAHAHILFRAGTVAPVCTTLLPCQTLLTDEEKRLIGQ 199
QY 277 EGYPIPTKLPLTKSEKALKIRKIKNKISAEKSRKKKMDSEKKBVSCSTENLEL 336
Db 200 EGVSLPSHLPLTKAEERVLKVRKIRNKQSAQDSRRKKEYIDGLSRVACSAQNOEL 259
QY 337 RKQVEVLENTNRTLLOQLQKLOTLVGMKVSRTCKLAGTGTCTCLMVVVLCPAVAFGSFFQ 396
Db 260 QKVQVELERHNISLVQLRQLQTL---IAQTSNKA-AQTSTCVLILFSLAILPSFS 314
QY 397 GYGPYPNATKWLPSQHPLEPYTASVVRSRNLIYEE 434
Db 315 PFQSRPEAG-----SEDYQPHGVTSRNLTHKD 342

Search completed: March 5, 2003, 06:18:33
Job time : 43 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 5, 2003, 06:20:31 ; Search time 356 Seconds
(without alignments)
61.715 Million cell updates/sec

Title: US-09-884-566A-2

Perfect score: 2650

Sequence: 1 MEVLESEQSVLQWDRKLS.....LEGNETLKVELERRVNA TP 521

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

188354 seqs, 42170167 residues

1 number of hits satisfying chosen parameters: 188354

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubaa/US08_NEW PUB pep.*
- 2: /cgn2_6/ptodata/1/pubaa/PCT_NEW PUB pep.*
- 3: /cgn2_6/ptodata/1/pubaa/US06_NEW PUB pep.*
- 4: /cgn2_6/ptodata/1/pubaa/US06_PUBCOMB pep.*
- 5: /cgn2_6/ptodata/1/pubaa/US07_NEW PUB pep.*
- 6: /cgn2_6/ptodata/1/pubaa/US07_PUBCOMB pep.*
- 7: /cgn2_6/ptodata/1/pubaa/PCTUS_PUBCOMB pep.*
- 8: /cgn2_6/ptodata/1/pubaa/US08_PUBCOMB pep.*
- 9: /cgn2_6/ptodata/1/pubaa/US09_NEW PUB pep.*
- 10: /cgn2_6/ptodata/1/pubaa/US09_PUBCOMB pep.*
- 11: /cgn2_6/ptodata/1/pubaa/US10_NEW PUB pep.*
- 12: /cgn2_6/ptodata/1/pubaa/US10_PUBCOMB pep.*
- 13: /cgn2_6/ptodata/1/pubaa/US60_NEW PUB pep.*
- 14: /cgn2_6/ptodata/1/pubaa/US60_PUBCOMB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2650	100.0	521	9	US-09-884-566-2
2	479	18.1	244	10	US-09-925-297-710
3	380	14.3	226	9	US-09-866-050A-627
4	358	13.5	252	10	US-09-764-864-1527
5	348	13.1	236	10	US-09-764-864-1526
6	348	13.1	239	10	US-09-764-864-1106
7	330.5	12.5	371	9	US-09-981-876-259
8	330.5	12.5	371	9	US-09-148-545-259
9	160.5	6.1	172	10	US-09-764-877-1459
10	160	6.0	44	10	US-09-864-761-46079
11	152	5.7	334	9	US-09-165-522-14
12	150	5.7	589	9	US-09-962-855-4
13	146.5	5.5	331	9	US-09-051-989-10
14	146.5	5.5	331	9	US-09-861-097-10
15	145.5	5.5	192	10	US-09-772-656-2
16	145.5	5.5	192	10	US-09-772-656-6
17	145.5	5.5	192	10	US-09-772-656-10
18	144.5	5.5	581	9	US-09-962-855-2
19	143.5	5.4	2005	10	US-09-735-367B-3

20	143.5	5.4	2063	10	US-09-735-367B-2	Sequence 2, Appli
21	142	5.4	230	9	US-10-050-185-27	Sequence 27, Appl
22	140.5	5.3	1404	10	US-09-811-045A-1	Sequence 1, Appli
23	139	5.2	1192	10	US-09-789-386-2	Sequence 2, Appli
24	139	5.2	1192	10	US-09-758-140-6	Sequence 6, Appli
25	139	5.2	1192	10	US-09-893-348-23	Sequence 23, Appli
26	139	5.2	1192	10	US-09-972-599A-6	Sequence 6, Appli
27	138	5.2	1781	10	US-09-738-877-3	Sequence 3, Appli
28	138	5.2	2462	9	US-09-819-104A-5	Sequence 5, Appli
29	134.5	5.1	941	12	US-10-124-557-14	Sequence 14, Appl
30	134.5	5.1	1022	12	US-10-124-557-84	Sequence 84, Appl
31	134.5	5.1	1038	12	US-10-124-557-74	Sequence 74, Appl
32	134.5	5.1	1049	12	US-10-124-557-58	Sequence 58, Appl
33	134.5	5.1	1140	12	US-10-124-557-104	Sequence 104, App
34	134.5	5.1	1270	12	US-10-124-557-44	Sequence 44, Appl
35	134.5	5.1	1311	12	US-10-124-557-42	Sequence 42, Appl
36	134.5	5.1	1313	12	US-10-124-557-142	Sequence 142, App
37	134.5	5.1	1314	12	US-10-124-557-50	Sequence 50, Appl
38	134.5	5.1	1320	12	US-10-124-557-46	Sequence 46, Appl
39	134.5	5.1	1320	12	US-10-124-557-60	Sequence 60, Appl
40	134.5	5.1	1354	12	US-10-124-557-48	Sequence 48, Appl
41	134.5	5.1	1361	12	US-10-124-557-40	Sequence 40, Appl
42	134.5	5.1	1363	12	US-10-124-557-52	Sequence 52, Appl
43	134.5	5.1	1404	12	US-10-124-557-2	Sequence 2, Appli
44	134.5	5.1	1404	12	US-10-124-557-62	Sequence 62, Appl
45	133	5.0	92	9	US-10-059-720-9	Sequence 9, Appli

ALIGNMENTS

RESULT 1
US-09-884-566-2
; Sequence 2, Application US/09884566
; Publication No. US2002019712A1
; GENERAL INFORMATION:
; APPLICANT: Keegan, Kathleen S. US2002019712A1el CREBa Isoform
; TITLE OF INVENTION: No. US2002019712A1el CREBa Isoform
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/884,566
; FILING DATE: 19-Jun-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/407,715
; FILING DATE: 28-Sep-1999
; APPLICATION NUMBER: US/09/005,970
; FILING DATE: 12-JAN-1998
; APPLICATION NUMBER: US/08/721,684
; FILING DATE: 27-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/33487
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 521 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2;
US-09-884-566-2

Query Match 100.0%; Score 2650; DB 9; Length 521;
Best Local Similarity 100.0%; Pred. No. 2.7e-168;
Matches 521; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEVLSEGEVSQVLDWRKLSLSEPGTEALMYTHFSELLDFSNQVQLQLSDPFLSK 60
DB 1 MEVLSEGEVSQVLDWRKLSLSEPGTEALMYTHFSELLDFSNQVQLQLSDPFLSK 60

QY 61 SSMSEVPSPAPLIQAEHSVSLSEEPRTQSPFFHAATSDSFNDEEVESEKWLSTEF 120
DB 61 SSMSEVPSPAPLIQAEHSVSLSEEPRTQSPFFHAATSDSFNDEEVESEKWLSTEF 120

QY 121 PSATIKKEPITEEQPPGLVPSVTLTITATSTPPEKEESPLDMNAGDSSCQTLPKIKLE 180
DB 121 PSATIKKEPITEEQPPGLVPSVTLTITATSTPPEKEESPLDMNAGDSSCQTLPKIKLE 180

QY 181 PHEVDQFLNFPKPEASVDQLHLPPTPPSHSDSEGLSPNPRLHPFSLQAHSPVRAMP 240
DB 181 PHEVDQFLNFPKPEASVDQLHLPPTPPSHSDSEGLSPNPRLHPFSLQAHSPVRAMP 240

QY 241 RGPALSTSPLLTAPHKLOGSGPLVLTBEKRTLVAEGYPIPTKLPTKSEKALKKIRR 300
DB 241 RGPALSTSPLLTAPHKLOGSGPLVLTBEKRTLVAEGYPIPTKLPTKSEKALKKIRR 300

QY 301 KIKNKISAESESRKKKYEYMDSEKVEKSCSTENLELRKKEVLENTNRTLQLOKLOTL 360
DB 301 KIKNKISAESESRKKKYEYMDSEKVEKSCSTENLELRKKEVLENTNRTLQLOKLOTL 360

QY 361 VMGKVRTCKLAGTGTGTCMLMVVLCFAVAFSGFGYGPYPATKMALPSQHPLSEPYT 420
DB 361 VMGKVRTCKLAGTGTGTCMLMVVLCFAVAFSGFGYGPYPATKMALPSQHPLSEPYT 420

QY 421 ASVVRNLLIYEHAPELSESSPASTGELGWDGSSLLRASSGLEALPEVDLPHPLIS 480
DB 421 ASVVRNLLIYEHAPELSESSPASTGELGWDGSSLLRASSGLEALPEVDLPHPLIS 480

QY 481 NETSLEKSVLLELQOHLVSKLSKGNETLKVLELERRVNTFF 521
DB 481 NETSLEKSVLLELQOHLVSKLSKGNETLKVLELERRVNTFF 521

RESULT 2
US-09-925-297-710
; Sequence 710, Application US/09925297
; Patent No. US20020081659A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA105
; CURRENT APPLICATION NUMBER: US/09/925,297
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05989
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 928
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 710
; LENGTH: 244
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (189)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (229)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-297-710

Query Match 18.1%; Score 479; DB 10; Length 244;
Best Local Similarity 57.4%; Pred. No. 1e-24;
Matches 105; Conservative 21; Mismatches 39; Indels 18; Gaps 3;

QY 280 PIPTKPLTKSEKALKKIRKKNKISAESESRKKKYEYMDSEKVEKSCSTENLELRKK 339
DB 1 PIPTKPLTKSEKALKKIRKKNKISAESESRKKKYEYMDSEKVEKSCSTENLELRKK 60

QY 340 VEVLENTNRTLQLOKLOTLVMGKVRTCKLAGTGTGTCMLMVVLCFAVAFSGFGYGP 399
DB 61 VETLENTNRTLQLOKLOTLVMGKVRTCKLAGTGTGTCMLMVVLCFAVAFSGFGYGP 120

QY 400 PYPATKMALPSQHPLSEPYTASVVRNLLIYEHAPELSESSPASTGELGWDGSS 457
DB 121 EFSSGSGTV--KEDPLADGVYASQMPSSRLFYDDGA-----GLWEDGR 164

QY 458 SLL 460
DB 165 STL 167

RESULT 3
US-09-866-050A-627
; Sequence 627, Application US/09866050A
; Publication No. US20030040471A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods for Their Use
; FILE REFERENCE: 11000.1011c4U
; CURRENT APPLICATION NUMBER: US/09/866,050A
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 725
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 627
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Rat
US-09-866-050A-627

Query Match 14.3%; Score 380; DB 9; Length 226;
Best Local Similarity 47.6%; Pred. No. 3.4e-18;
Matches 90; Conservative 18; Mismatches 55; Indels 26; Gaps 3;

QY 300 RKIKNKISAESESRKKKYEYMDSEKVEKSCSTENLELRKKEVLENTNRTLQLOKLOTL 359
DB 1 RKIKNKISAESESRKKKYEYMDSEKVEKSCSTENLELRKKEVLENTNRTLQLOKLOTL 60

QY 360 LVMGKVRTCKLAGTGTGTCMLMVVLCFAVAFSGFGYGPYPATKMALPSQHPLSEPY 419
DB 61 LVTSKISRYPYKMAATGTGTCMLMVVLCFAVAFSGFGYGPYPATKMALPSQHPLSEPY 120

QY 420 TASVVRNLLIYEHAPELSESSPASTGELGWDGSSLLRASSGLEALPEVDLPHPLIS 467
DB 121 AASQMPSSRLFYDDGA-----GSWEDGHRGALLPVEPPEGWELKPGGPAAE 166

QY 468 ALPEVDLPH 476
DB 167 PRPDHLRH 175

RESULT 4
US-09-764-864-1527
; Sequence 1527, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.

;; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;; FILE REFERENCE: PT223
;; CURRENT APPLICATION NUMBER: US/09/764,864
;; CURRENT FILING DATE: 2001-01-17
;; Prior application data removed - consult PALM or file wrapper
;; NUMBER OF SEQ ID NOS: 1792
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 1527
;; LENGTH: 252
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: SITE
;; LOCATION: (188)
;; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-864-1527

Query Match 13.5%; Score 358; DB 10; Length 252;
Best Local Similarity 36.5%; Pred. No. 1.1e-16;
Matches 95; Conservative 42; Mismatches 81; Indels 42; Gaps 6;

Qy 221 NPLHPSLSQHSVPRAMPGPSALSTSPLLTAPHKLGSGPLVLTDEEKRTLVAEGYP 280
Db 10 NSRVDPRVRAHILPRAGTVAPVCTT-----LLPCQTLFLTDEEKRLLGQGV 60
Qy 281 IPTKLPLTKSEKALKTKIRKIKNISQESRRKKKEYMDSLEKKVSCSTENLELRKV 340
Db 61 LPSHLPLTKAEERVLKVRKIRNKQSAQDSRRKKKEYIDGLESRVAACSAQNELQKV 120
Qy 341 EVLENTNTLQLOKLOTLVMGKVSRTCKLAGTGTGCLMVLVLCFAVAFSGFQGY 400
Db 121 QELERHNSLVLAQLQQLT-----AQTSTCVLILFSLALILPSPFQS 175
Qy 401 YPSATKMALPSQHPLEPYTASVVRNLLIYEE-----HAPLEE-----SS 442
Db 176 RPEAG-----SEDYQPHXVTSRNILTHKDVTENLTQVVSRLREPPGAXDANGS 225
Qy 443 SPASTGELGGWDGSSLR 462
Db 226 TRTLLKMGKPRPSGRIS 245

RESULT 5
US-09-764-864-1526
; Sequence 1526, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1526
; LENGTH: 236
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-864-1526

Query Match 13.1%; Score 348; DB 10; Length 236;
Best Local Similarity 45.0%; Pred. No. 4.8e-16;
Matches 77; Conservative 33; Mismatches 46; Indels 15; Gaps 3;

Qy 264 LVLTEEEKRTLVAEGYPPTKPLTKSEKALKTKIRKIKNISQESRRKKKEYMDSLE 323
Db 28 LFLTDEEKRLLGQGVSPSHLPLTKAEERVLKVRKIRNKQSAQDSRRKKKEYIDGLE 87
Qy 324 KKVESCSTENLELRKKEVLENTNTLQLOKLOTLVMGKVSRTCKLAGTGTGCLMV 383
Db 88 SRVAACSAQNELQKQVLEHRHNSLVLAQLRQLT-----IAQTSNKA-AQTSTCVLIL 142

Qy 384 VLCFAVAFSGFQGYPPYPATKMLPSQHPLEPYTASVVRNLLIYEE 434
Db 143 LFSALALILPSPFPQSRPEAG-----SEDYQPHRVTSRNILTHKD 183

RESULT 6
US-09-764-864-1106
; Sequence 1106, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1106
; LENGTH: 239
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (225)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-864-1106

Query Match 13.1%; Score 348; DB 10; Length 239;
Best Local Similarity 45.0%; Pred. No. 4.9e-16;
Matches 77; Conservative 33; Mismatches 46; Indels 15; Gaps 3;

Qy 264 LVLTEEEKRTLVAEGYPPTKPLTKSEKALKTKIRKIKNISQESRRKKKEYMDSLE 323
Db 31 LFLTDEEKRLLGQGVSPSHLPLTKAEERVLKVRKIRNKQSAQDSRRKKKEYIDGLE 90
Qy 324 KKVESCSTENLELRKKEVLENTNTLQLOKLOTLVMGKVSRTCKLAGTGTGCLMV 383
Db 91 SRVAACSAQNELQKQVLEHRHNSLVLAQLRQLT-----IAQTSNKA-AQTSTCVLIL 145
Qy 384 VLCFAVAFSGFQGYPPYPATKMLPSQHPLEPYTASVVRNLLIYEE 434
Db 146 LFSALALILPSPFPQSRPEAG-----SEDYQPHRVTSRNILTHKD 186

RESULT 7
US-09-981-876-259
; Sequence 259, Application US/09981876
; Patent No. US20020164669A1
; GENERAL INFORMATION:
; TITLE OF INVENTION: 70 Human Secreted Proteins
; FILE REFERENCE: P2001P1
; CURRENT APPLICATION NUMBER: US/09/981,876
; CURRENT FILING DATE: 2001-10-19
; PRIOR APPLICATION NUMBER: 09/148,545
; PRIOR FILING DATE: 1998-09-04
; PRIOR APPLICATION NUMBER: 60/040,162
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,333
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/038,621
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,161
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,626
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,334
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,336
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,163
; PRIOR FILING DATE: 1997-03-07

[illegible]

```
; PRIOR APPLICATION NUMBER: 60/056,632
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,664
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,876
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,881
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,909
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,875
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,862
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,887
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,908
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/048,964
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/057,650
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: 60/056,884
; PRIOR FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 280
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 259
; LENGTH: 371

Query Match          12.5%  Score 330.5; DB 9; Length 371;
Best Local Similarity 30.7%; Pred. No. 1.3e-14;
Matches 116; Conservative 56; Mismatches 123; Indels 83; Gaps 15;

Qy 150 STPEKESPLDMAGDSQCTLIPIKLEPHEVDQFLNFPKESASVDQLHLP-PTPPS 208
Db 25 TAPDEAVRAPLDW-----ALP-----LSEVPSDWEVDLLCSLLSPPA 62

Qy 209 SHSDSEGLSPNPL--HPFSLQAHSPVRAMPGRGALSTSPLLTAPH-----KQSGS 262
Db 63 SLNLTLS-----SSNPLCVHHDHTYSLPRETVSMDLSESCRKEGTQMTQHMELAEQIEA 118

Qy 263 PLVLTTEERKTLVARGYPIPTKLPTKSEKALKIRKKNKISAOESRRKKKEYMDSL 322
Db 119 RLVLTEBSKLSLEKGLLIPETLPTKTEEQLEKRRKRRKRNKRAQSSRRKKVYVGGL 178

Qy 323 EKKVSCSTENLELRKKEVLENTNRTLLQQLKQLTLMGVKSVRTCKLAGTGTCTCLMV 382
Db 179 ESRVLKYTAQNWELQNLQLEEQNLSDLDQLRKLQANWVIEISNKT-----SSSSTCLIV 233

Qy 383 VVLCFAVAFGFFQGYGYGPSATKVALPSQH-----PLSPYTSVSVRSRNLIIY 432
Db 234 LLVSFCLLLVP-----AMYSSTRGSLPAEHGVLRSQRLALPSEDYPQLEL----- 279

Qy 433 EEHAPLESSSPA-STGELGWDGSS--LLRASSGLEAL-----PEVDLPHELI 479
Db 280 ----PALQSEVPKDSHTQ---WLDGSDCVLQAPGNTSCLLHYMPQASPELPWFPPDL 332

Qy 480 SNETSLEKSVLLEQOHL 497
Db 333 SSE-PLCRGPILPLQANL 349

RESULT 8
US-09-148-545-259
; Sequence 259, Application US/09148545
; Publication No. US20030027132A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 70 Human Secreted Proteins
; FILE REFERENCE: P2001P1
; CURRENT APPLICATION NUMBER: US/09/148,545
; CURRENT FILING DATE: 1998-09-04
; EARLIER APPLICATION NUMBER: PCT/US98/04482
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,671
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,568
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,580
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/047,601
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,632
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,612
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,582
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,596
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,581
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,584
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,500
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,587
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,492
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,598
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,613
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,582
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,596
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,612
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,632
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,601
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,580
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,568
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,314
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,569
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,311
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,671
; EARLIER FILING DATE: 1997-04-11
```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 5, 2003, 06:17:30 ; Search time 43 Seconds
(without alignments)
356.496 Million cell updates/sec

Title: US-09-884-566A-2
Perfect score: 2650
Sequence: 1 MEVLESGSQSVLQWRKLSLSEPCETALMYHVFSELLDFNSQNLGQLLSDPFLSEK 521

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues
al number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5A.COMB.pap.*
2: /cgn2_6/ptodata/1/iaa/5B.COMB.pap.*
3: /cgn2_6/ptodata/1/iaa/6A.COMB.pap.*
4: /cgn2_6/ptodata/1/iaa/6B.COMB.pap.*
5: /cgn2_6/ptodata/1/iaa/PCTUS.COMB.pap.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES									
Result No.	Score	Query		Length	DB	ID	Description		
		Match							
1	2650	100.0		521	2	US-08-721-684C-2	Sequence 2, Appli		
2	2850	100.0		521	2	US-09-005-970-2	Sequence 2, Appli		
3	2850	100.0		521	4	US-09-407-715-2	Sequence 2, Appli		
4	395	14.9		395	4	US-09-247-155-113	Sequence 113, App		
5	146.5	5.5		327	2	US-07-684-965-6	Sequence 6, Appli		
6	146.5	5.5		327	4	US-09-215-098-5	Sequence 5, Appli		
7	146.5	5.5		327	4	US-09-252-658-6	Sequence 6, Appli		
8	146.5	5.5		331	1	US-08-094-533B-10	Sequence 10, Appli		
9	146.5	5.5		331	1	US-08-276-860A-10	Sequence 10, Appli		
10	146.5	5.5		331	1	US-08-444-393-10	Sequence 10, Appl		
11	146.5	5.5		331	1	US-08-799-913-10	Sequence 10, Appl		
12	146.5	5.5		331	2	US-08-711-893-10	Sequence 10, Appl		
13	146.5	5.5		331	2	US-09-150-200-10	Sequence 10, Appl		
14	146.5	5.5		331	3	US-09-150-201-10	Sequence 10, Appl		
15	146.5	5.5		331	4	US-09-452-370-10	Sequence 10, Appl		
16	146.5	5.5		331	4	US-09-461-649-10	Sequence 10, Appl		
17	146.5	5.5		331	5	PCT-US95-15353-39	Sequence 39, Appl		
18	144.5	5.5		392	4	US-09-416-050A-2	Sequence 2, Appli		
19	144.5	5.5		392	4	US-09-664-800-2	Sequence 2, Appli		
20	144.5	5.5		392	4	US-09-665-309-2	Sequence 2, Appli		
21	144.5	5.5		392	4	US-09-661-569-2	Sequence 2, Appli		
22	143.5	5.4		295	2	US-08-343-443B-107	Sequence 107, App		
23	142	5.4		230	4	US-09-387-418A-27	Sequence 27, Appl		
24	140.5	5.3		1404	4	US-08-801-308-1	Sequence 1, Appli		
25	137.5	5.2		331	5	PCT-US94-08119-10	Sequence 10, Appl		
26	137.5	5.2		331	5	PCT-US94-08120-10	Sequence 10, Appl		
27	137.5	5.2		331	5	PCT-US94-12913A-10	Sequence 10, Appl		

28	134.5	5.1	941	4	US-07-757-022B-14	Sequence 14, Appl
29	134.5	5.1	1022	4	US-07-757-022B-84	Sequence 84, Appl
30	134.5	5.1	1038	4	US-07-757-022B-74	Sequence 74, Appl
31	134.5	5.1	1049	4	US-07-757-022B-58	Sequence 58, Appl
32	134.5	5.1	1140	4	US-07-757-022B-104	Sequence 104, Appl
33	134.5	5.1	1270	4	US-07-757-022B-44	Sequence 44, Appl
34	134.5	5.1	1311	4	US-07-757-022B-42	Sequence 42, Appl
35	134.5	5.1	1313	4	US-07-757-022B-142	Sequence 142, Appl
36	134.5	5.1	1314	4	US-07-757-022B-50	Sequence 50, Appl
37	134.5	5.1	1320	4	US-07-757-022B-46	Sequence 46, Appl
38	134.5	5.1	1320	4	US-07-757-022B-60	Sequence 60, Appl
39	134.5	5.1	1354	4	US-07-757-022B-48	Sequence 48, Appl
40	134.5	5.1	1361	4	US-07-757-022B-40	Sequence 40, Appl
41	134.5	5.1	1363	4	US-07-757-022B-52	Sequence 52, Appl
42	134.5	5.1	1404	4	US-07-757-022B-2	Sequence 2, Appli
43	134.5	5.1	1404	4	US-07-757-022B-62	Sequence 62, Appli
44	133	5.0	92	4	US-09-299-495F-9	Sequence 9, Appli
45	133	5.0	93	2	US-08-690-011A-9	Sequence 9, Appli

ALIGNMENTS

RESULT 1
US-08-721-684C-2
; Sequence 2, Application US/08721684C
; Patent No. 5854016
; GENERAL INFORMATION:
; APPLICANT: Keegan, Kathleen S.
; TITLE OF INVENTION: No. 5854016el CREBa Isoform
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/721.684C
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/33487
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 521 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-721-684C-2

Query Match 100.0%; Score 2650; DB 2; Length 521;
Best Local Similarity 100.0%; Pred. No. 5.8e-215;
Matches 521; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MEVLESGSQSVLQWRKLSLSEPCETALMYHVFSELLDFNSQNLGQLLSDPFLSEK 60		
Db	1	MEVLESGSQSVLQWRKLSLSEPCETALMYHVFSELLDFNSQNLGQLLSDPFLSEK 60		
Qy	61	SESMEVEPSPTSPAPLIAEHHSYLSSEPRQSPPTHAATSDSFNDEEVESEKWLSTEF 120		
Db	61	SESMEVEPSPTSPAPLIAEHHSYLSSEPRQSPPTHAATSDSFNDEEVESEKWLSTEF 120		

Best Local Similarity 100.0%; Pred. No. 5.8e-215; Matches 521; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 PSATIKKEPITEEQPPGLVPSVTLTITATISTPFEKEESPLDMNAGDSSCQTLIPKIKLE 180
DB 121 PSATIKKEPITEEQPPGLVPSVTLTITATISTPFEKEESPLDMNAGDSSCQTLIPKIKLE 180
QY 181 PHEVQDFLNFSKPEASVDQLHLPPTPPSSHSDSGSLSPNRLHPFSLQAHSVPVRAMP 240
DB 181 PHEVQDFLNFSKPEASVDQLHLPPTPPSSHSDSGSLSPNRLHPFSLQAHSVPVRAMP 240
QY 241 RGPSALSTSPLLTAPHKLGSGPLVTEBEKRTLVAEGYPIPTKPLTKSEKALKKIRR 300
DB 241 RGPSALSTSPLLTAPHKLGSGPLVTEBEKRTLVAEGYPIPTKPLTKSEKALKKIRR 300
QY 301 KIKNKISAQESRRKKKEYMDSLEKKVESCSTENLELRKKVEVLENTNRTLLOQLKQTL 360
DB 301 KIKNKISAQESRRKKKEYMDSLEKKVESCSTENLELRKKVEVLENTNRTLLOQLKQTL 360
QY 361 VMKVSRTCKLAGTGTCTLMVVVLCFAVAFSGFFQGYGYPYSATKMALPSSHPLSEPT 420
DB 361 VMKVSRTCKLAGTGTCTLMVVVLCFAVAFSGFFQGYGYPYSATKMALPSSHPLSEPT 420
QY 421 ASVVRSRNLLIYEEHAPLEBESSPASTGBLGWDGSSLLRASSGLEALPEVDLPHFLIS 480
DB 421 ASVVRSRNLLIYEEHAPLEBESSPASTGBLGWDGSSLLRASSGLEALPEVDLPHFLIS 480
QY 481 NETSLEKSVLLELQHLVSSKLEGNETLKVLELRRVNTF 521
DB 481 NETSLEKSVLLELQHLVSSKLEGNETLKVLELRRVNTF 521

RESULT 2
US-09-005-970-2
; Sequence 2, Application US/09005970
; Patent No. 5959079
; GENERAL INFORMATION:
; APPLICANT: Keegan, Kathleen S.
; TITLE OF INVENTION: No. 5959079el CREBa Isoform
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/721,684
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/33487
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 521 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-005-970-2

Query Match . 100.0%; Score 2650; DB 2; Length 521;

TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 521 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-407-715-2

Query Match 100.0%; Score 2650; DB 4; Length 521;
Best Local Similarity 100.0%; Pred. No. 5.8e-215;
Matches 521; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MEVLSEGSQVLDNRKLSLSEPGETALMYHFSSELLDEFSQVNLGQLLSDPFLSEK 60
Db 1 MEVLSEGSQVLDNRKLSLSEPGETALMYHFSSELLDEFSQVNLGQLLSDPFLSEK 60
61 SESMEVEPSPTAPLIAEHSYSLSEPRQSPETHAATSDSFNDEEVESEKWLSTEF 120
Db 61 SESMEVEPSPTAPLIAEHSYSLSEPRQSPETHAATSDSFNDEEVESEKWLSTEF 120
Qy 121 PSATIKKEPITEEOPPGVLVPSVTLTITAIPTPFKEESPLDMNAGDSSCOTLIPIKILE 180
Db 121 PSATIKKEPITEEOPPGVLVPSVTLTITAIPTPFKEESPLDMNAGDSSCOTLIPIKILE 180
Qy 181 PHEVDQFLNFPKESASVDQLHLPTPPSSHSSDSGSLSPNRLHPFSLQAHSVPVRAMP 240
Db 181 PHEVDQFLNFPKESASVDQLHLPTPPSSHSSDSGSLSPNRLHPFSLQAHSVPVRAMP 240
Qy 241 RPSALSTSPLLTAPHKLOGGVLVLTBEEKRTLVAEGYPIPTKPLTKSEBKALKIKRR 300
Db 241 RPSALSTSPLLTAPHKLOGGVLVLTBEEKRTLVAEGYPIPTKPLTKSEBKALKIKRR 300
Qy 301 KIKNKISAEQSRKKKYMDSLEKKVESCSTENLELRKKVELENTRTLQOLQKLOTL 360
Db 301 KIKNKISAEQSRKKKYMDSLEKKVESCSTENLELRKKVELENTRTLQOLQKLOTL 360
Qy 361 VMGKVSRTCKLAGTGTGTCMLVMVVLCFVAFSGFPQGYGPSATKMLPSQHPLSEPYT 420
Db 361 VMGKVSRTCKLAGTGTGTCMLVMVVLCFVAFSGFPQGYGPSATKMLPSQHPLSEPYT 420
Qy 421 ASVVRNRLLIYEEHAPLEESSSPASTGELGWDGSGSLLRASSGLEALPEVDLPHFLIS 480
Db 421 ASVVRNRLLIYEEHAPLEESSSPASTGELGWDGSGSLLRASSGLEALPEVDLPHFLIS 480
481 NETSLEKSVLLEQHLVSSKLEGNETLKVVELERRVNATF 521
Db 481 NETSLEKSVLLEQHLVSSKLEGNETLKVVELERRVNATF 521

RESULT 4
US-09-247-155-113
Sequence 113, Application US/09247155A
Patent No. 6312922
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, Jean-Baptiste
APPLICANT: Duclert, Aymeric
APPLICANT: Bougueleret, Lydie
TITLE OF INVENTION: Complementary DNAs
FILE REFERENCE: GENSET.021A
CURRENT APPLICATION NUMBER: US/09/247.155A
CURRENT FILING DATE: 1999-02-09
EARLIER APPLICATION NUMBER: 60/074.121
EARLIER FILING DATE: 1998-02-09
EARLIER APPLICATION NUMBER: 60/081.563
EARLIER FILING DATE: 1998-04-13
EARLIER APPLICATION NUMBER: 60/096.116
EARLIER FILING DATE: 1998-08-10
EARLIER APPLICATION NUMBER: 60/099.273
EARLIER FILING DATE: 1998-10-04
NUMBER OF SEQ ID NOS: 182

SOFTWARE: Patent.pm
SEQ ID NO 113
LENGTH: 395
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SIGNAL
LOCATION: -310..-1
US-09-247-155-113

Query Match 14.9%; Score 395; DB 4; Length 395;
Best Local Similarity 34.0%; Pred. No. 2.8e-25;
Matches 115; Conservative 49; Mismatches 108; Indels 66; Gaps 12;
Qy 135 PPGLVPSVTLTITAIPTPFKEESPLDMNAGDSSC--QTLP-----KIKLEPHV--DQ 186
Db 33 PPPEVPVTRL-----QEQLGQWKGSGDRCGLQESEPEDFLKLFIDPNEVYCS 82
Qy 187 FLNFPKESASVDQLHLPTPPSSHSSDS-----EGSLSPNRLHPFSL 230
Db 83 ASPGSDSGISEDSCH-PDPPAPRATSPMLYEVVEAGALERMQGETGPNVGLISIQD 141
Qy 231 QAHSVPVRAMPGPSALSTSPLLTAPHKLOGSG-----PLVLTTEEEXRTLVA 276
Db 142 Q-WSPAFNVP-DSCMVSELFPDAHAHILPRAGTVAPVCTTLLPCQTLFLTDEEKELCQ 199
Qy 277 EGYPIPTKPLTKSEKALKIKRKIRKIKNKISAEQSRKKKYMDSLEKKVESCSTENLEL 336
Db 200 EGVSLPSHLPTKAEERVLKVRKIRNKQAQDSRRRKEYIDGLSRVAACSAQNOEL 259
Qy 337 RKKVEVLENTRTLQOLQKLOTLVMGKVSRTCKLAGTGTGTCMLVMVVLCFVAFSGFPQ 396
Db 260 QKKVQELERHNLISVAQRLOTL----IAQTSNKA-AQTCTCVLILLFLSLAILPLPSFS 314
Qy 397 GYGVPVPSATKMLPSQHPLSEPYTASVVRNRLLIYEE 434
Db 315 PFQSRPEAG-----SEDIQPHGVTSRNLTHTKD 342

RESULT 5
US-07-684-965-6
Sequence 6, Application US/07684965
Patent No. 5919649
GENERAL INFORMATION:
APPLICANT: HABENER, JOEL F.
APPLICANT: HOEFELER, JAMES P.
TITLE OF INVENTION: A-CAMP RESPONSIVE TRANSCRIPTIONAL
TITLE OF INVENTION: ENHANCER BINDING PROTEIN
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1225 Connecticut Avenue, N.W., Suite 300
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/684.965
FILING DATE: 19910522
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Cimbala, Michele A
REGISTRATION NUMBER: 33,851
REFERENCE/DOCKET NUMBER: 0609.2470004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)833-7533
TELEFAX: (202)833-8716
INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:
LENGTH: 327 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-684-965-6

Query Match 5.5%; Score 146.5; DB 2; Length 327;
Best Local Similarity 23.8%; Pred. No. 0.00018;
Matches 73; Conservative 52; Mismatches 105; Indels 77; Gaps 15;

QY 74 APLIAQHSYSLSEEPRTQSPFTHAATSDSFNDEEVESEK-WY-----LSTFPFSATIK 126

DB 80 SQVQTVQISTIAEHSQESVD--SVTDSQNRRLSRPSRYKILNDLSDAFQVPI 137

QY 127 KSPITEQPPGLVPSVTITAITSTPFKEESP--LDMNAGDSSCQTLIPKIKLEPHEV 184

DB 138 EEEKSEETS--APAIT-TVT-VPTPIYQTSQGVIATIQGG-----AIQLANNGT 184

QY 185 DQFLNFPKPEASVDQLHLPPTPS-----SHSDSEGLSPNPRHLHPSLSQAHPVPA 238

DB 185 DGVQGLQ-----TLTWTNAAATPGTILQYAGTTDQQLVPTNQV---VVOAASGDVQT 237

QY 239 MPRGPSALSTPLLT-APHKLOGSPLVLTBEEKRTLVAEGYPIPTKPLTKSEKALKK 297

DB 238 Y-----QIKTAPTSTIAPGVNMASSP-----ALPTQAEAPRKR 272

QY 298 IRRKINKISAQESRRKKKYMDSLEKKVESCSTENLELRKKVLENTNRTLQLOQLK 357

DB 273 EVRLMKNREARCRKKKKEYVKCLE-----NRVAVLENQNKTLIELKAL 318

QY 358 QTLVMGK 364

DB 319 KOLYCHK 325

RESULT 6

US-09-215-098-5

Sequence 5, Application US/09215098

Patent No. 6194632

GENERAL INFORMATION:

APPLICANT: Leiden, Jeffery M

TITLE OF INVENTION: DILATED CARDIOMYOPATHY IN TRANSGENIC MICE EXPRESSING A

TITLE OF INVENTION: DOMINANT-NEGATIVE CREB TRANSCRIPTION FACTOR IN THE

FILE REFERENCE: 9189-4

CURRENT APPLICATION NUMBER: US/09/215,098

PRIOR FILING DATE: 1998-12-18

PRIOR APPLICATION NUMBER: 60/068,011

PRIOR FILING DATE: 1997-12-18

NUMBER OF SEQ ID NOS: 5

SOFTWARE: Patent In Ver. 2.1

SEQ ID NO 5

LENGTH: 327

TYPE: PRT

ORGANISM: Homo sapiens

US-09-215-098-5

Query Match 5.5%; Score 146.5; DB 4; Length 327;
Best Local Similarity 22.6%; Pred. No. 0.00018;
Matches 85; Conservative 56; Mismatches 114; Indels 121; Gaps 18;

QY 50 QLLSDPPLSE-KSESMEVEPSP-----TSPAP-----L 76

DB 10 QQSGDAATEAENQOMTVQAPQIATLAQVMPAAHATSSAPTTLVQLPNGQTVQVHV 69

QY 77 IQAHSYSLSEEPRTQ-----SPFTHAATSDSFNDEEVESEK-WY-----LS 117

DB 70 IQAQP-SVQSPQVQTVQISTIAESDSQESVDSVTDQKRRILRRRSYRKILNDLS 128

QY 118 TEPPSATIKKEPTEPPGLVPSVTITAITSTPFKEESP--LDMNAGDSSCQTLIP 175

DB 129 SDAPGVRIEERESEETS--APAIT-TVT-VPTPIYQTSQGVIATIQGG----- 175

QY 176 KIKLEPHEVDQFLNFPKPEASVDQLHLPPTPS-----SHSDSEGLSPNPRHLHPSL 229

DB 176 AIQLANNGTDGVQGLQ-----TLTWTNAAATPGTILQYAGTTDQQLVPSNQV---VV 228

QY 230 SQAHSPPVRAMPGRPSALSTPLLT-APHKLOGSPLVLTBEEKRTLVAEGYPIPTKPL 288

DB 229 QAASGDVQTY-----QIKTAPTSTIAPGVNMASSP-----ALPTQ 263

QY 289 KSEKALKIKIRKINKISAQESRRKKKYMDSLEKKVESCSTENLELRKKVLENTN 348

DB 264 PAEEAARREVRLMKNREARCRKKKKEYVKCLE-----NRVAVLENQNK 309

QY 349 TILQLOQLKQTLVMGK 364

DB 310 TLIBELKALKOLYCHK 325

RESULT 7

US-09-252-658-6

Sequence 6, Application US/09252658

Patent No. 6251667

GENERAL INFORMATION:

APPLICANT: HABENER, JOEL P.

APPLICANT: HOFFLER, JAMES P.

TITLE OF INVENTION: A CAMP-RESPONSIVE TRANSCRIPTIONAL

TITLE OF INVENTION: ENHANCER BINDING PROTEIN

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

ADDRESS: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.

STREET: 1100 NEW YORK AVE., NW, STE. 600

CITY: WASHINGTON

STATE: DC

COUNTRY: USA

ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/252,658

FILING DATE: HERewith

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/684,965

FILING DATE: 22-MAY-1991

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US89/05234

FILING DATE: 20-NOV-1989

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/272,980

FILING DATE: 18-NOV-1988

ATTORNEY/AGENT INFORMATION:

NAME: Cimballa, Michele A

REGISTRATION NUMBER: 33,851

REFERENCE/DOCKET NUMBER: 0609.1990002

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202)371-2600

TELEFAX: (202)371-2540

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 327 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-252-658-6

Query Match 5.5%; Score 146.5; DB 4; Length 327;

Best Local Similarity 23.8%; Pred. No. 0.00018;

Matches 73; Conservative 52; Mismatches 105; Indels 77; Gaps 15;

QY 74 APLIAQHSYSLSEEPRTQSPFTHAATSDSFNDEEVESEK-WY-----LSTFPFSATIK 126

Db 80 SPQVQVQISIAESESQESVD--SVTDSQNRRLRRYSRKILNDLSADGCVPR1 137
Qy 127 KEPITEOPPGVLPSVTLTITAIPTPFKEBSP--LDMNAGDSSCQTLIPKIKLEPHEV 184
Db 138 EEKSEETS--APAIT-TVT-VPTPIYQTSQGVIAITQGG-----AIQLANNGT 184
Qy 185 DQFLNFSPEASVDOLHLPPTPPS-----SHSSDSEGLSPNPLHPLFSLSOAHSVPYR 238
Db 185 DGVOGLQ-----TLTWNAAATOPGTTILOAYQTTDGOQILVPTNQV---VVOAASGDVQT 237
Qy 239 MPRGPSALSTPLLT-APHKLGSGPLVLTBEEKRTLVAEGVPIPTKLPLTKSEBALKK 297
Db 238 Y-----QIRTAFTSTIAGVWASSP-----ALPTQPAEEAPKR 272
Qy 298 IRRKIKISAOESRRKKKYNDLSLEKKVESCSTENLELRKKVEYLENTRTLLOQL 357
Db 273 EVRLMKNREARECRKKKYVKCLE-----NRVALENQNTLIELKAL 318
358 QTLVWGK 364
319 KOLYCHK 325

RESULT 8

US-08-094-533B-10
; Sequence 10, Application US/08094533B
; Patent No. 5534426
; GENERAL INFORMATION:
; APPLICANT: KARIN, MICHAEL
; APPLICANT: MASAHICO, HIBI
; APPLICANT: ANNING, LIN
; TITLE OF INVENTION: ONCOPROTEIN PROTEIN KINASE
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FISH & RICHARDSON P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: California
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/094,533B
; FILING DATE: 19-JUL-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Ph.D., Lisa A.,
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07257/006001 (PD2848)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 678-5070
; TELEFAX: (619) 678-5099
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 331 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-094-533B-10

Query Match 5.5%; Score 146.5; DB 1; Length 331;

Best Local Similarity 20.3%; Pred. No. 0.00018;

Matches 70; Conservative 60; Mismatches 141; Indels 73; Gaps 10;

Qy 68 PSPTSP-----APLIQAEHSYSLSEEPRTQSPFFTHAATSDSFNDEEV-----ESK 114
Db 20 PSESGPYGYNPKILKQSMNTLNADPVGSLKPLHRAKNSDLLTSPDVGLLKLASPELERL 79

Qy 115 YL-----STEPSATIKKEPITEOPPGVLPSVTLTITAIPTPFKEBSPDLMNAGG 166
Db 80 IQSSNGHITTTPTQFCLCPKNVTDEQ-EGFAEGFVRALAEHLS----- 123
Qy 167 DSSCQTLIPKIKLEPHEVDQFLNFSPEASV-----DQLHLPTPPSSHSSDSEG 216
Db 124 ---QNTLPSVTSAAQPVNGAGWAPAVASVAGSGSGGFSASLHSEPPVYANLSNPNP 179
Qy 217 SLSNPNRHPFSLSOAHSVPVAMPGRGPSALSTPLLTAPHKLQSGGP-----LVLTEBEK 271
Db 180 ALSGGGAPSYGAAGL-----AFPAQPOQQO---QPPHLPQPMVQHPRLQALKEEP 229
Qy 272 RTLVAEGVPIPTKLPLTKSEBALKKIRKIKISAOESRRKKKYNDLSLEKKVESCST 331
Db 230 QTVPEMGETPPLPIDMESQERIKAEKRMNRNIAASKCRKRLERLARLEEKVKTLKA 289
Qy 332 ENLELRKKVEYLENTRTLLOQLKQTLVNGKVSRTCKLAGTQ 375
Db 290 QNSE-----LASTANWLREOVAQLKQKVMNVHNSGCOLMLTQ 326

RESULT 9

US-08-276-860A-10
; Sequence 10, Application US/08276860A
; Patent No. 5593884
; GENERAL INFORMATION:
; APPLICANT: KARIN, MICHAEL
; APPLICANT: MASAHICO, HIBI
; APPLICANT: ANNING, LIN
; TITLE OF INVENTION: ONCOPROTEIN PROTEIN KINASE
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FISH & RICHARDSON P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: California
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/276,860A
; FILING DATE: 18-JUL-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Ph.D., Lisa A.,
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07257/017001 (PD3701)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 678-5070
; TELEFAX: (619) 678-5099
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 331 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-276-860A-10

Query Match 5.5%; Score 146.5; DB 1; Length 331;

Best Local Similarity 20.3%; Pred. No. 0.00018;

Matches 70; Conservative 60; Mismatches 141; Indels 73; Gaps 10;

Qy 68 PSPTSP-----APLIQAEHSYSLSEEPRTQSPFFTHAATSDSFNDEEV-----ESK 114
Db 20 PSESGPYGYNPKILKQSMNTLNADPVGSLKPLHRAKNSDLLTSPDVGLLKLASPELERL 79
Qy 115 YL-----STEPSATIKKEPITEOPPGVLPSVTLTITAIPTPFKEBSPDLMNAGG 166
Db 80 IQSSNGHITTTPTQFCLCPKNVTDEQ-EGFAEGFVRALAEHLS----- 123

Db 124 ----QNTLPSVTSAAQPVNGAGMVAPAVASVAGSGGGSFASLSHSBPPVYANLSNPNP 179
Qy 217 SLSPNRLHPFSLQASHSPVRAMPRGSPALSTPLTAPHKLQSGP-----LVLTEEEK 271
Db 180 ALSGGGAPSAGAGL-----APPAQPOQQO-----QPPHLPQOMPVOHPRLOALKEEP 229
Qy 272 RTLVAEGVPIPTKLPLTKSEBALKIRKIKNKISAOESRRKKKKEYMDSLEKKVESCST 331
Db 230 QTVEMPGETPLSPIDMESQERIKAEKRMNRNIAASKCKRKLRIARLEEKVKTLKA 289
Qy 332 ENLELRKKVELENTNRTLLQLOKLQTLVMGKVSRTCKLAGTQ 375
Db 290 QNSE-----LASTANMLREQVAQLKQKVMNHVNSGQQLMLTQ 326

RESULT 11
US-08-799-913-10
; Sequence 10, Application US/08799913
; Patent No. 5804399
; GENERAL INFORMATION:
; APPLICANT: KARIN, MICHAEL
; APPLICANT: HIBI, MASAHIKO
; APPLICANT: LIN, ANNING
; TITLE OF INVENTION: ONCOPROTEIN PROTEIN KINASE
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FISH & RICHARDSON P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: California
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/799,913
; FILING DATE: 13-FEB-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/444,393
; FILING DATE: 19-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Ph.D., Lisa A.,
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07257/017002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 678-5070
; TELEFAX: (619) 678-5099
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 331 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-799-913-10

Query Match 5.5%; Score 146.5; DB 1; Length 331;
Best Local Similarity 20.3%; Pred. No. 0.00018;
Matches 70; Conservative 60; Mismatches 141; Indels 73; Gaps 10;
Qy 68 PSPTSP-----APLIQAEHSYLSSEPRTPSPHTAATSDSFNDEEV-----ESEKW 114
Db 20 PSSEGPYGNPKILKQSMTNLADPVGSLKPHLRKNSDLLTSPDVGLLKLASPELERL 79
Qy 115 YL-----STFPSSATIKKEPITERQPPGLVPSVTLTITATSTPPEKESPLDMNAGG 166
Db 80 IQSSNGHIHTTPTTQFLCPKNVTDEQ-EGFAGFVRALAEHLS----- 123
Qy 167 DSSCQTLIPKILKEPHEVDQFLNFSPEASV-----DQLHLPTTPSSHSSDSEG 216

Qy 167 DSSCQTLIPKILKEPHEVDQFLNFSPEASV-----DQLHLPTTPSSHSSDSEG 216
Db 124 ----QNTLPSVTSAAQPVNGAGMVAPAVASVAGSGGGSFASLSHSBPPVYANLSNPNP 179
Qy 217 SLSPNRLHPFSLQASHSPVRAMPRGSPALSTPLTAPHKLQSGP-----LVLTEEEK 271
Db 180 ALSGGGAPSAGAGL-----APPAQPOQQO-----QPPHLPQOMPVOHPRLOALKEEP 229
Qy 272 RTLVAEGVPIPTKLPLTKSEBALKIRKIKNKISAOESRRKKKKEYMDSLEKKVESCST 331
Db 230 QTVEMPGETPLSPIDMESQERIKAEKRMNRNIAASKCKRKLRIARLEEKVKTLKA 289
Qy 332 ENLELRKKVELENTNRTLLQLOKLQTLVMGKVSRTCKLAGTQ 375
Db 290 QNSE-----LASTANMLREQVAQLKQKVMNHVNSGQQLMLTQ 326

RESULT 10
US-08-444-393-10
; Sequence 10, Application US/08444393
; Patent No. 5605808
; GENERAL INFORMATION:
; APPLICANT: KARIN, MICHAEL
; APPLICANT: HIBI, MASAHIKO
; APPLICANT: LIN, ANNING
; TITLE OF INVENTION: ONCOPROTEIN PROTEIN KINASE
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FISH & RICHARDSON P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: California
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/444,393
; FILING DATE: 19-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Ph.D., Lisa A.,
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07257/017002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 678-5070
; TELEFAX: (619) 678-5099
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 331 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-444-393-10

Query Match 5.5%; Score 146.5; DB 1; Length 331;
Best Local Similarity 20.3%; Pred. No. 0.00018;
Matches 70; Conservative 60; Mismatches 141; Indels 73; Gaps 10;
Qy 68 PSPTSP-----APLIQAEHSYLSSEPRTPSPHTAATSDSFNDEEV-----ESEKW 114
Db 20 PSSEGPYGNPKILKQSMTNLADPVGSLKPHLRKNSDLLTSPDVGLLKLASPELERL 79
Qy 115 YL-----STFPSSATIKKEPITERQPPGLVPSVTLTITATSTPPEKESPLDMNAGG 166
Db 80 IQSSNGHIHTTPTTQFLCPKNVTDEQ-EGFAGFVRALAEHLS----- 123
Qy 167 DSSCQTLIPKILKEPHEVDQFLNFSPEASV-----DQLHLPTTPSSHSSDSEG 216

Db 124 ---QNTLPSTVTSAAQPVNGAGWAPAVASVAGSGSGFSASLHSEPPVYANLSNFNPG 179
Qy 217 SLSPNRLHPHLSQAHSFVPRAMPGRPSALSTPLTAPHKLGQSGP-----LVITEBEK 271
Db 180 ALSGGGAPSYGAAGL-----AFPAQPOQQ-----QPPHLLPQOMPVOHPRLQALKEEP 229
Qy 272 RTLVAEGYPIPTKLPLTKSEKALKIRKKNKISAOESRRKKKEYMDSLEKKYVESCST 331
Db 230 QTVPEMPGETPLSPIDMESQERIKAEKRMNRNIAASKCRKKRLERTARLEEKVTKLKA 289
Qy 332 ENLELRKKVEVLNTRTLLOQLQKLTLMVGKVSRTCKLAGTQ 375
Db 290 QNSE-----LASTANMLREQVAQLKQKVMNVHNSGCOLMLTQ 326

RESULT 12

US-08-711-893-10
; Sequence 10, Application US/08711893
; Patent No. 5837244
; GENERAL INFORMATION:
; APPLICANT: KARIN, MICHAEL
; APPLICANT: MASAHICO, HIBI
; APPLICANT: ANNING, LIN
; TITLE OF INVENTION: ONCOPROTEIN PROTEIN KINASE
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FISH & RICHARDSON P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: California
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/711,893
; FILING DATE: 12-SEP-1996
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/276,860
; FILING DATE: 18-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Ph.D., Lisa A.,
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07257/017001 (PD3701)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 678-5070
; TELEFAX: (619) 678-5099
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 331 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-711-893-10

Query Match 5.5%; Score 146.5; DB 2; Length 331;
Best Local Similarity 20.3%; Pred. No. 0.00018;
Matches 70; Conservative 60; Mismatches 141; Indels 73; Gaps 10;

Qy 68 PSPTSP-----APLIQAHSYSLSEEPRTQSPFFTHAATSDFNDEEV-----ESEKW 114
Db 20 PSESGPYGYNPKILKQSMNTLNADPVGSLKPLHRAKNSDLTSPDVGLLKLASFELERL 79
Qy 115 YL-----STEPSATIKKEPTEEPQGLVPSVTITITAITPFEKESPLDMNAGG 166
Db 80 IIQSSNGHITTTPTPTQPLCPKNVTDEQ-EGFAEGFVRALAEHLS----- 123
Qy 167 DSSCQTLIPKILKLPHEVDQFLNFSPEASV-----DQLHLPTPPSSHSDSEG 216

Db 124 ---QNTLPSTVTSAAQPVNGAGWAPAVASVAGSGSGFSASLHSEPPVYANLSNFNPG 179
Qy 217 SLSPNRLHPHLSQAHSFVPRAMPGRPSALSTPLTAPHKLGQSGP-----LVITEBEK 271
Db 180 ALSGGGAPSYGAAGL-----AFPAQPOQQ-----QPPHLLPQOMPVOHPRLQALKEEP 229
Qy 272 RTLVAEGYPIPTKLPLTKSEKALKIRKKNKISAOESRRKKKEYMDSLEKKYVESCST 331
Db 230 QTVPEMPGETPLSPIDMESQERIKAEKRMNRNIAASKCRKKRLERTARLEEKVTKLKA 289
Qy 332 ENLELRKKVEVLNTRTLLOQLQKLTLMVGKVSRTCKLAGTQ 375
Db 290 QNSE-----LASTANMLREQVAQLKQKVMNVHNSGCOLMLTQ 326

RESULT 13

US-09-150-200-10
; Sequence 10, Application US/09150200
; Patent No. 5994513
; GENERAL INFORMATION:
; APPLICANT: KARIN, MICHAEL
; APPLICANT: HIBI, MASAHICO
; APPLICANT: LIN, ANNING
; TITLE OF INVENTION: ONCOPROTEIN PROTEIN KINASE
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FISH & RICHARDSON P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: California
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/150,200
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/444,393
; FILING DATE: 19-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Ph.D., Lisa A.,
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07257/017002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 678-5070
; TELEFAX: (619) 678-5099
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 331 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-150-200-10

Query Match 5.5%; Score 146.5; DB 2; Length 331;
Best Local Similarity 20.3%; Pred. No. 0.00018;
Matches 70; Conservative 60; Mismatches 141; Indels 73; Gaps 10;

Qy 68 PSPTSP-----APLIQAHSYSLSEEPRTQSPFFTHAATSDFNDEEV-----ESEKW 114
Db 20 PSESGPYGYNPKILKQSMNTLNADPVGSLKPLHRAKNSDLTSPDVGLLKLASFELERL 79
Qy 115 YL-----STEPSATIKKEPTEEPQGLVPSVTITITAITPFEKESPLDMNAGG 166
Db 80 IIQSSNGHITTTPTPTQPLCPKNVTDEQ-EGFAEGFVRALAEHLS----- 123
Qy 167 DSSCQTLIPKILKLPHEVDQFLNFSPEASV-----DQLHLPTPPSSHSDSEG 216

Db 124 ----QNTLPVSTSAAPVNGAGMVAFAVASVAGSGGGSFASLSHSEPPVYANISNFNP 179
Qy 217 SLSPNRLHPFSLQAHSPVRAMPGRGPSALSTSLTAPHKLGSGP-----LVUTEEEK 271
Db 180 ALSSGGGAPSAGAGL-----APPAQPOQQO-----QPPHLPQOMPVOHPRLQALKEEP 229
Qy 272 RTLVAEGVPIPTKLPLTKSEKALKIRKIKNKISAQESRRKKKEYMDSLEKKVESCST 331
Db 230 QTVPEMGETPPLSPIDMESQERIKAEKRMNRNIAASKCKRKLRLARLEEKVKTLKA 289
Qy 332 ENLELRKKVELVNTNRTLQLOKLOQLTVMGKVSRTCKLAGTQ 375
Db 290 QNSE-----LASTANMLREQVAQLKOKVMNVHNSGCOLMLTQ 326

RESULT 14
US-09-150-201-10
; Sequence 10, Application US/09150201
; Patent No. 6001584
; GENERAL INFORMATION:
; APPLICANT: KARIN, MICHAEL
; APPLICANT: HIBI, MASAHIKO
; APPLICANT: LIN, ANNING
; TITLE OF INVENTION: ONCOPROTEIN PROTEIN KINASE
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FISH & RICHARDSON P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: California
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/150,201
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/444,393
; FILING DATE: 19-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Ph.D., Lisa A.,
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07257/017002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 678-5070
; TELEFAX: (619) 678-5099
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 331 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-150-201-10

Query Match 5.5%; Score 146.5; DB 3; Length 331;
Best Local Similarity 20.3%; Pred. No. 0.00018;
Matches 70; Conservative 60; Mismatches 141; Indels 73; Gaps 10;
Qy 68 PSPTSP-----APLQAEHSYSLESEPTQSPFTHAATSDSFNDEEV-----ESEKW 114
Db 20 PSSEGPYGNPKLIKOSMTNLADPVGSLKPHLRKNSDLTSPDVGLLKLASPELERL 79
Qy 115 YL-----STEPSATIKKEPITEBPQGLVPSVTLTITTAISTPPEKESPLDMNAGG 166
Db 80 IQSSNGHITTTPTTQFLCPKNVTDEQ-EGFAGGFVRLAELHS----- 123
Qy 167 DSSCQTLIPKLEFHEVDQFLNFSPEASV-----DQLHLPTTPSSHSSDSEG 216

Db 124 ----QNTLPVSTSAAPVNGAGMVAFAVASVAGSGGGSFASLSHSEPPVYANISNFNP 179
Qy 217 SLSPNRLHPFSLQAHSPVRAMPGRGPSALSTSLTAPHKLGSGP-----LVUTEEEK 271
Db 180 ALSSGGGAPSAGAGL-----APPAQPOQQO-----QPPHLPQOMPVOHPRLQALKEEP 229
Qy 272 RTLVAEGVPIPTKLPLTKSEKALKIRKIKNKISAQESRRKKKEYMDSLEKKVESCST 331
Db 230 QTVPEMGETPPLSPIDMESQERIKAEKRMNRNIAASKCKRKLRLARLEEKVKTLKA 289
Qy 332 ENLELRKKVELVNTNRTLQLOKLOQLTVMGKVSRTCKLAGTQ 375
Db 290 QNSE-----LASTANMLREQVAQLKOKVMNVHNSGCOLMLTQ 326
RESULT 15
US-09-452-370-10
; Sequence 10, Application US/09452370
; Patent No. 6193965
; GENERAL INFORMATION:
; APPLICANT: KARIN, MICHAEL
; APPLICANT: HIBI, MASAHIKO
; APPLICANT: LIN, ANNING
; TITLE OF INVENTION: ONCOPROTEIN PROTEIN KINASE
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FISH & RICHARDSON P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: California
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/452,370
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/150,200
; FILING DATE:
; APPLICATION NUMBER: US 08/444,393
; FILING DATE: 19-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Ph.D., Lisa A.,
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07257/017002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 678-5070
; TELEFAX: (619) 678-5099
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 331 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-452-370-10

Query Match 5.5%; Score 146.5; DB 4; Length 331;
Best Local Similarity 20.3%; Pred. No. 0.00018;
Matches 70; Conservative 60; Mismatches 141; Indels 73; Gaps 10;
Qy 68 PSPTSP-----APLQAEHSYSLESEPTQSPFTHAATSDSFNDEEV-----ESEKW 114
Db 20 PSSEGPYGNPKLIKOSMTNLADPVGSLKPHLRKNSDLTSPDVGLLKLASPELERL 79
Qy 115 YL-----STEPSATIKKEPITEBPQGLVPSVTLTITTAISTPPEKESPLDMNAGG 166
Db 80 IQSSNGHITTTPTTQFLCPKNVTDEQ-EGFAGGFVRLAELHS----- 123

```
QY 167 SSSQTLIPKIKLEPHEVDQFLNFSPEASV-----DQLHLPTTPSSHSSDSEG 216
Db 124 ----QNTLPSVTISAAQPVNGAGMVAVASVAGSGSGGFSASLHSEPPVYANLSNPNPG 179
QY 217 SLSNPRLHPFSLSOAHSPVRAMPSPGSAISTPLLTAPHKLGSGP-----LVLTEEEK 271
Db 180 ALSSGGGAPSYGAAGL-----AFPAPQOQO-----QPPHLPQOMPVOHPRLQALKEEP 229
QY 272 RTLVAEGYPIPTKPLTKSEBKALKIRKIKNKISAQESRRKKKKEYMDSLEKKVESCST 331
Db 230 QTVPEMPCETPPLSPIDWESQERIKAEKRMNRNRTAASKCRKRKLERTARLEEKVKTLLKA 289
QY 332 ENLELRKKVEVLENTNRTLLOLOKLOTLVMGKVSRTCKLAGTQ 375
Db 290 QNSE-----LASTANMLREQVAQLKOKVNVHNVNSGCCQLMLTQ 326
```

Search completed: March 5, 2003, 06:21:35
Job time : 64 secs

THIS PAGE BLANK (USPTO)

This Page is inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☒ BLACK BORDERS
- ☒ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☒ FADED TEXT OR DRAWING
- ☐ BLURED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☒ COLORED OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REPERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images
problems checked, please do not report the
problems to the IFW Image Problem Mailbox**

THIS PAGE BLANK (USPTO)